# ACTA MATHEMATICA

## ACADEMIAE SCIENTIARUM HUNGARICAE



REDIGIT

G. HAJÓS

TOMUS VI FASCICULI 1—2

MAGYAR TUDOMÁNYOS AKADÉMIA
BUDAPEST, 1955

ACTA MATH. HUNG.

# ACTA MATHEMATICA
## ACADEMIAE SCIENTIARUM HUNGARICAE

A MAGYAR TUDOMÁNYOS AKADÉMIA III. OSZTÁLYÁNAK
MATEMATIKAI KÖZLEMÉNYEI

SZERKESZTŐSÉG ÉS KIADÓHIVATAL: BUDAPEST, V., ALKOTMÁNY U. 21


Az Acta Mathematica orosz, francia, angol és német nyelven közöl értekezéseket a matematika köréből.

Az Acta Mathematica változó terjedelmű füzetekben jelenik meg, több füzet alkot egy kötetet.

A közlésre szánt kéziratok géppel írva, a következő címre küldendők:

*Acta Mathematica, Budapest 62, Postafiók 440.*

Ugyanerre a címre küldendő minden szerkesztőségi és kiadóhivatali levelezés.

Az Acta Mathematica előfizetési ára kötetenként belföldre 80 forint, külföldre 110 forint. Megrendelhető a belföld számára az „Akadémiai Kiadó"-nál (Budapest, V., Alkotmány utca 21. Bankszámla 04-878-111-46), a külföld számára pedig a „Kultúra" Könyv- és Hírlap Külkereskedelmi Vállalatnál (Budapest, VI., Sztálin út 21. Bankszámla 43-790-057-181) vagy külföldi képviseleteinél és bizományosainál.

---

*„Acta Mathematica"* публикует трактаты из области математических наук на русском, французском, английском и немецком языках.

*„Acta Mathematica"* выходит отдельными выпусками разного объема. Несколько выпусков составляют один том.

Предназначенные для публикации рукописи (в напечатанном на машинке виде) следует направлять по адресу:

*Acta Mathematica*, Budapest 62, Postafiók 440.

По этому же адресу направлять всякую корреспонденцию для редакции и администрации.

Подписная цена *„Acta Mathematica"* — 110 форинтов за том. Заказы принимает предприятие по внешней торговле книг и газет *„Kultúra"* (*Budapest, VI., Sztálin út 21.* Текущий счет № 43-790-057-181), или его заграничные представительства и уполномоченные.

# ÜBER DIE KONVERGENZ DER ORTHOGONALPOLYNOMENTWICKLUNGEN

Von

G. ALEXITS (Budapest), Mitglied der Akademie

**1.** Vor kurzem haben wir gezeigt,[1] daß die Fourierreihe einer Funktion $f \in L^2$ im Intervall $(a, b) \subset (0, 2\pi)$ fast überall konvergiert, wenn der quadratische Stetigkeitsmodul

$$\omega_2(\delta, f; a, b) = \sup_{|h| \leqq \delta} \left\{ \int_a^b [f(x+h) - f(x)]^2 dx \right\}^{\frac{1}{2}}$$

von der Größenordnung $O(1/\sqrt{\lambda(1/\delta)})$ ist, wo $\lambda(x)$ eine beliebige monoton zunehmende Funktion mit

$$\int_1^\infty \frac{dx}{x\lambda(x)} < \infty \tag{1}$$

bedeutet. Ungefähr zur selben Zeit erschien eine Arbeit von S. B. STEČKIN,[2] in welcher derselbe Satz in etwas verschiedener Fassung bewiesen wurde. In der vorliegenden Note setzen wir uns den Beweis einer ähnlichen Konvergenzbedingung bezüglich der Orthogonalpolynomentwicklungen stetiger Funktionen zum Ziel.

**2.** Bezeichne $w(x) \geqq 0$ eine in $(-1, 1)$ definierte, $L$-integrierbare Funktion. Sie bestimmt bekanntlich eindeutig ein System $\{p_n(x)\}$ von normierten Orthogonalpolynomen, wo $p_n(x)$ ein Polynom $n$-ten Grades bedeutet. Die Orthogonalreihe

$$\sum_{n=0}^\infty c_n p_n(x) \sim f(x) \tag{2}$$

[1] G. ALEXITS, Über den Einfluß der Struktur einer Funktion auf die Konvergenz fast überall ihrer Fourierreihe, *Acta Math. Acad. Sci. Hung.*, 4 (1953), S. 95–101.

[2] С. Б. Стечкин, О теореме Колмогорова—Селиверстова, Известия Акад. Наук СССР, 17 (1953), S. 499—512.

heißt die Entwicklung von $f(x)$, wenn ihre Koeffizienten durch

$$c_n = \int_{-1}^{1} w(x) f(x) p_n(x) dx$$

bestimmt sind. Wir werden den folgenden Satz beweisen:

*Genügt die Gewichtsfunktion $w(x)$ in $(-1, 1)$ fast überall der Bedingung*

$$0 < w(x) = O\left(\frac{1}{\sqrt{1-x^2}}\right)$$

*und sind die durch $w(x)$ bestimmten Orthonormalpolynome $p_0(x), p_1(x), \ldots$ in einem ganz im Inneren von $(-1, 1)$ liegenden Intervall $(a, b)$ in ihrer Gesamtheit beschränkt, erfüllt ferner der in $(a, b)$ gebildete Stetigkeitsmodul*

$$\omega(\delta, f; a, b) = \sup_{\substack{|h| \leqq \delta \\ a \leqq x \leqq b}} |f(x+h) - f(x)|$$

*der $L_w^2$-integrierbaren Funktion $f(x)$ die Bedingung*

$$\omega(\delta, f; a, b) = O\left(\frac{1}{\sqrt{\lambda(1/\delta)}}\right),$$

*wo $\lambda(x)$ eine monoton zunehmende, der Bedingung* (1) *genügende Funktion bedeutet, so konvergiert die Entwicklung* (2) *in $(a, b)$ fast überall.*

Bezeichnet insbesondere $\{p_n(x)\}$ eines der klassischen Orthogonalpolynomsysteme, so folgt aus unserem Satz als Korollar die Behauptung, daß die klassischen Orthogonalpolynomentwicklungen einer Funktion, die einer Dini-Lipschitzschen Bedingung $\alpha$-ter Ordnung mit $\alpha > \frac{1}{2}$ genügt, fast überall konvergieren. Dann läßt sich nämlich $(a, b)$ im Inneren von $(-1, 1)$ beliebig annehmen und $\lambda(x) = (\log x)^{2\alpha}$ mit $\alpha > \frac{1}{2}$ wählen.

**3.** Zum Beweise bilden wir $(-1, 1)$ durch die Transformation $\vartheta = \arccos x$ auf das Intervall $(0, \pi)$ ab, und betrachten statt $f(x)$ die gerade Funktion $f^*(\vartheta) = f(\cos \vartheta)$. Ist $|\chi| \leqq \delta$, so ist auch $|\cos(\vartheta + \chi) - \cos \vartheta| \leqq |\chi| \leqq \delta$ und daher gilt

$$\omega(\delta, f^*; \arccos b, \arccos a) \leqq \omega(\delta, f; a, b) = O\left(\frac{1}{\sqrt{\lambda(1/\delta)}}\right).$$

Da $(\arccos b, \arccos a)$ ganz im Inneren von $(0, \pi)$ liegt, läßt sich eine Funktion $g^*(\vartheta)$ folgenderweise definieren:

$$g^*(\vartheta) = \begin{cases} f^*(\vartheta), \text{ falls } \vartheta \in (\arccos b, \arccos a), \\ f^*(\arccos \mathrm{b}), \text{ falls } \vartheta \in (0, \arccos b), \\ f^*(\arccos a), \text{ falls } \vartheta \in (\arccos a, \pi). \end{cases}$$

Offensichtlich ist dann

(3) $$\omega(\delta, g^*; 0, \pi) = \omega(\delta, f^*; \arccos b, \arccos a) = O\left(\frac{1}{\sqrt{\lambda(1/\delta)}}\right).$$

Dann gilt aber — wie wir es in der eingangs erwähnten Note gezeigt haben[3] — für die Fourierkoeffizienten $\alpha_0, \alpha_1, \ldots$ der Funktion $g^*(\vartheta)$ die Kolmogoroff-Seliverstoff-Plessnersche Beziehung

$$\sum_{n=1}^{\infty} \alpha_n^2 (\log n)^2 < \infty.$$

An anderer Stelle haben wir aber bewiesen,[4] daß diese Bedingung, unter den getroffenen Annahmen über die Gewichtsfunktion $w(x)$, die Konvergenz der Entwicklung

$$\sum_{n=0}^{\infty} \gamma_n p_n(x) \sim g(x)$$

in $(-1, 1)$ fast überall nach sich zieht, wo $g(x)$ die aus $g^*(\vartheta)$ durch die Transformation $x = \cos\vartheta$ hervorgehende Funktion bezeichnet. Damit wäre unser Satz schon bewiesen, wenn wir wüßten, daß die Reihen $\Sigma c_n p_n(x)$ und $\Sigma \gamma_n p_n(x)$ in jedem Intervall $(a+\varepsilon, b-\varepsilon)$ mit beliebig kleinem $\varepsilon > 0$ äquikonvergent sind. Dies ergibt sich aber leicht. Die Partialsummen der beiden Entwicklungen stehen nämlich nach der Christoffel—Darbouxschen Formel in folgendem Zusammenhang:

$$s_n(f, x) - s_n(g, x) = O(1) \int_{-1}^{1} [f(t) - g(t)]\, w(t) \frac{p_n(t)\, p_{n+1}(x) - p_{n+1}(t) p_n(x)}{t-x}\, dt.$$

Da aber für $\vartheta \in (\arccos b, \arccos a)$ die Beziehung $f^*(\vartheta) = g^*(\vartheta)$ besteht, ist $f(t) - g(t) = 0$, wenn $t \in (a, b)$, also

$$s_n(f, x) - s_n(g, x) = O(1)\left(\int_{-1}^{a} + \int_{b}^{1}\right) \frac{f(t)-g(t)}{t-x}\, w(t) [p_n(t) p_{n+1}(x) - p_{n+1}(t) p_n(x)]\, dt.$$

Hier besteht die rechte Seite aus Entwicklungskoeffizienten der Funktion

$$F(t) = \begin{cases} 0 \text{ für } t \in (a, b), \\ \dfrac{f(t)-g(t)}{t-x} \text{ für } t \notin (a, b), \end{cases}$$

multipliziert mit den Polynomen $p_{n+1}(x)$ bzw. $p_n(x)$, deren absolute Werte aber, wegen $x \in (a+\varepsilon, b-\varepsilon)$, laut Annahme unter einer festen Zahl bleiben.

[3] A. a. O.[1], S. 98—99.

[4] G. Alexits, Sur la convergence et la sommabilité presque partout des séries de polynomes orthogonaux, *Acta Sci. Math.*, 12 (1950), S. 223—225.

Daraus folgt, da $F(t)$ eine $L_w^2$-integrierbare Funktion ist, deren Entwicklungskoeffizienten mithin gegen Null konvergieren:

$$s_n(f, x) - s_n(g, x) = o(1).$$

Hiemit ist die Äquikonvergenz der Reihen $s_n(f, x)$ und $s_n(g, x)$ im Intervall $(a+\varepsilon, b-\varepsilon)$ und folglich auch unser Satz bewiesen.

*(Eingegangen am 12. November 1954.)*

## О СХОДИМОСТИ РАЗЛОЖЕНИЯ В РЯД ПО ОРТОГОНАЛЬНЫМ МНОГОЧЛЕНАМ

Г. Алексич (Будапешт)

(Резюме)

Пусть $w(x)$ функция определенная в интервале $(-1, 1)$ и удовлетворяющая условию

$$0 < w(x) = O\left(\frac{1}{\sqrt{1-x^2}}\right).$$

Обозначим через $\{p_n(x)\}$ полную ортогональную нормированную систему полиномов определенную весовой функцией $w(x)$. Обозначим далее через $\omega(\delta, f;\, a, b)$ модуль непрерывности функции $f(x)$ в интервале $(a, b) \subset (-1, 1)$.

В настоящей работе доказывается аналог для ортогональных полиномов теоремы о рядах Фурье, доказанной независимо друг от друга Стечкиным и автором.

Если $\{p_n(x)\}$ является равномерно ограниченной в интервале $(a, b)$ лежащего внутри $(-1, 1)$, и модуль непрерывности функции $f(x)$ удовлетворяет условию

$$\omega(\delta, f;\, a, b) = O\left(\frac{1}{\sqrt{\lambda(1/\delta)}}\right),$$

где $\lambda(x)$ функция, монотонно возрастающая для которой

$$\int_1^\infty \frac{dx}{x\lambda(x)} < \infty,$$

то разложение

$$f(x) \sim \sum_{n=0}^\infty c_n p_n(x)$$

сходится почти всюду в $(a, b)$.

# ZETAFUNKTIONEN IN DER ALGEBRA

Von

LADISLAUS RÉDEI (Szeged), Mitglied der Akademie

*Dem Andenken von* Bernhard Riemann *gewidmet anläßlich der Hundertjahrfeier des Einreichens seiner Habilitationsschrift an die Göttinger Universität*

## § 1. Einleitung

Auch schon bisher sind viele wichtige Verallgemeinerungen und Analoga der Riemannschen Zetafunktionen bekannt geworden. In dieser Arbeit werden wir im Zusammenhang mit beliebigen (algebraischen) Strukturen (d. h. Gruppen, Ringen usw.) eine neue Art „Zetafunktionen" einführen. Näher werden wir hier nur den Fall von endlichen Abelschen Gruppen untersuchen, der zu Resultaten führen wird, die uns von großem Interesse und fortsetzungsfähig zu sein scheinen. Vielversprechend ist auch der Fall der unendlichen Abelschen Gruppen (insbesondere der Torsionsgruppen), den wir aber bisher noch kaum untersucht haben. Die Nützlichkeit unserer Zetafunktionen für sonstige Strukturen scheint uns dagegen fraglich zu sein. Selbst schon die endlichen Abelschen Gruppen werden zu ziemlich zusammengesetzten Untersuchungen Anlaß geben. Am merkwürdigsten wird dabei vielleicht, daß alle von uns bezüglich der Zetafunktionen für die endlichen Abelschen Gruppen entdeckten Grundtatsachen vom Fundamentalsatz dieser Gruppen unabhängig sind. Diese Unabhängigkeit erschwert einerseits die Untersuchungen, gibt ihnen aber andererseits auch den Reiz.

Für eine endliche Menge $M$ bezeichnen wir mit $(M)$ die Anzahl der Elemente von $M$. Ist $M$ die leere Menge, die wir mit $\mathbf{0}$ bezeichnen, so soll natürlich $(M)=0$ sein. Für eine unendliche Menge $M$ setzen wir $(M)=\infty$. Ist $M$ insbesondere eine Struktur, so nennen wir $(M)$ die Ordnung von $M$.

Stets soll $n$ eine nichtnegative ganze Zahl und $\mathfrak{N}$ die Menge der Zahlen $1,\ldots,n$ bezeichnen. (Für $n=0$ ist also $\mathfrak{N}=\mathbf{0}$.)

Für eine komplexe Variabel $z$ bezeichnen wir mit $Rz$ den reellen Teil von $z$.

Sind $A_1, \ldots, A_n$ irgendwelche (nicht notwendig verschiedene) Unterstrukturen[1] einer beliebigen algebraischen Struktur $S$, so setzen wir für $Rz > 0$

$$(1) \qquad \varrho(z) = \varrho(z; A_1, \ldots, A_n) = \sum_{\mathfrak{M} \subseteq \mathfrak{N}} (-1)^{(\mathfrak{M})} (A_{\mathfrak{M}})^{-z},$$

wobei wir mit $A_{\mathfrak{M}}$ die durch die Elemente der $A_i$ $(i \in \mathfrak{M})$ erzeugte Unterstruktur von $S$ bezeichnen, d. h. $A_{\mathfrak{M}}$ ist die engste Unterstruktur von $S$, die die $A_i$ $(i \in \mathfrak{M})$ enthält. (Wenn $\mathfrak{M} = \mathbf{0}$ ist, so kann $A_{\mathfrak{M}}$ sinnlos werden, aber dann setzen wir $(A_{\mathfrak{M}}) = 1$.) Man sieht, daß (1) aus $2^n$ Gliedern besteht. Gleich bemerken wir, daß die $\mathfrak{M}$ mit $(\mathfrak{M}) = \infty$ wegen $Rz > 0$ keinen Beitrag zu (1) liefern, weshalb man in (1) nur die $\mathfrak{M}$ mit endlichem $A_{\mathfrak{M}}$ zu berücksichtigen hat.

Wir lassen auch den Fall zu, daß $S$ mit Operatoren versehen ist; dann sollen als „Unterstrukturen“ nur die „zulässigen Unterstrukturen“ gelten.

Handelt es sich um (abzählbar) unendlich viele Unterstrukturen $A_1, A_2, \ldots$ von $S$, so definieren wir für ein $z$ mit $Rz \geqq 1$

$$(2) \qquad \varrho(z) = \varrho(z; A_1, A_2, \ldots) = \lim_{n \to \infty} \varrho(z; A_1, \ldots, A_n)$$

stets, wenn die rechte Seite existiert und von der Reihenfolge der $A_i$ unabhängig ist. Ist dabei $\varrho(z)$ in einem Teilgebiet der Halbebene $Rz \geqq 1$ analytisch und regulär, so definieren wir $\varrho(z)$ (von diesem Gebiet ausgehend) durch analytische Fortsetzung.

Als ein sehr einfaches Beispiel betrachten wir den Fall, wo $S$ eine Gruppe und zwar das direkte Produkt von $A_1, A_2, \ldots$ bedeutet und $(A_i) = p_i$ die $i$-te Primzahl de. natürlichen Zahlenfolge ist. Nach (1) gilt dann offenbar

$$\varrho(z; A_1, \ldots, A_n) = \prod_{i=1}^{n} (1 - p_i^{-z}).$$

Hieraus folgt nach (2)

$$\varrho(z) = \varrho(z; A_1, A_2, \ldots) = \prod_{i=1}^{\infty} (1 - p_i^{-z}).$$

Jetzt ist also $\varrho(z)^{-1}$ eben die Zetafunktion von RIEMANN.

Deshalb fühlen wir uns veranlaßt ganz allgemein die Bezeichnung

$$(3) \qquad \zeta(z) = \zeta(z; A_1, A_2, \ldots) = \varrho(z)^{-1} = \varrho(z; A_1, A_2, \ldots)^{-1}$$

einzuführen und diese $\zeta(z)$ *Zetafunktionen* (in der Algebra) zu nennen. Je nachdem die vorgelegte Struktur eine Gruppe, ein Ring, ... ist, sprechen wir entsprechend auch über *gruppentheoretische, ringtheoretische, ... Zetafunktionen.* (In der Definition (3) lassen wir auch den Fall zu, wo an Stelle von $A_1, A_2, \ldots$ eine endliche Folge $A_1, \ldots, A_n$ tritt.)

[1] „Unterstrukturen“ einer Gruppe, eines Ringes usw. sollen bzw. die Untergruppen Unterringe usw. der gegebenen Struktur bedeuten.

Ein Fall für die ringtheoretischen Zetafunktionen entsteht z. B. folgendermaßen. Es bezeichne $S$ einen Ring, den wir als den Modul $S^+$ seiner Elemente auffassen, versehen mit den Elementen von $S$ als links- und rechtsseitigen Operatoren, wobei als „Operation" die links- bzw. rechtsseitige Multiplikation in $S$ gelten soll. Das kommt darauf hinaus, daß in (3) für die $A_i$ eben nur die Ideale von $S$ zuzulassen sind und $A_{\mathfrak{M}}$ (in (1)) die Summe der Ideale $A_i$ $(i \in \mathfrak{M})$ bedeutet. Beachtet man nur die linksseitigen Operationen, so hat man für die $A_i$ die Linksideale von $S$ zuzulassen. Im „gewöhnlichen" ringtheoretischen Fall ist dagegen ein Ring $S$ ohne Operatoren angegeben; für die $A_i$ sind dann die Unterringe von $S$ zuzulassen, ferner bezeichnet $A_{\mathfrak{M}}$ jetzt den durch die Elemente der $A_i$ $(i \in \mathfrak{M})$ erzeugten Unterring von $S$.

Unsere Definitionen lassen sich für den gruppen- und ringtheoretischen Fall auch „dualisieren". Das soll einfach darin bestehen, daß wir in (1) $A_{\mathfrak{M}}$ als den Durchschnitt der $A_i$ $(i \in \mathfrak{M})$ deuten und $(A_{\mathfrak{M}})$ durch den Index $(S : A_{\mathfrak{M}})$ von $A_{\mathfrak{M}}$ in $S$ ersetzen. Die Erweiterung dieses „dualen" Falles für unendlich viele $A_i, A_2, \ldots$ soll ähnlich wie bei (2) durch Grenzübergang und analytische Fortsetzung geschehen. Die so entstehenden $\zeta(z)$ nennen wir ebenfalls Zetafunktionen, genauer *duale Zetafunktionen* (in der Algebra).

Leicht sieht man, daß die Dedekindschen Zetafunktionen der algebraischen Zahlentheorie als sehr einfacher Spezialfall hierher gehören. Man bezeichne nämlich mit $S$ den Ring der ganzen Zahlen eines absolut algebraischen Zahlkörpers vom endlichen Grade und nehme für $A_1, A_2, \ldots$ die sämtlichen Primideale von $S$. Für die entsprechende (duale) Zetafunktion gilt dann offenbar

$$\zeta(z; A_1, \ldots, A_n) = \prod_{i=1}^{n} (1 - N(A_i)^{-z})^{-1},$$

wobei $N(A_i)$ den Index $(S : A_i)$, d. h. die Norm von $A_i$ bezeichnet. Hiernach ist jetzt $\zeta(z; A_1, A_2, \ldots)$ in der Tat mit der Dedekindschen Zetafunktion identisch.

In dieser Arbeit wollen wir, wie gesagt, näher nur die gruppentheoretischen Zetafunktionen für die endlichen Abelschen Gruppen (kurz den endlich-Abelschen Fall) untersuchen. In diesem Fall sind die zwei Arten Zetafunktionen nichtverschieden, denn die anfangs definierten Zetafunktionen angewendet auf die Charaktergruppe der vorgelegten Gruppe stimmen mit den dualen Zetafunktionen für die letztere Gruppe überein, wie man das leicht sieht. Im unendlich Abelschen Falle ist es natürlich anders, aber wir bemerken, daß die Riemannsche Zetafunktion auch als Spezialfall unter den dualen Zetafunktionen für die unendlichen Abelschen Gruppen enthalten ist. Sind nämlich $A_1, A_2, \ldots$ die sämtlichen Untergruppen von Primzahlindex in der unendlichen zyklischen Gruppe, so stimmt die duale Zetafunktion $\zeta(z; A_1, A_2, \ldots)$ mit der eben betrachteten Dedekindschen Zetafunktion für den Fall des rationalen Zahlkörpers, d. h. mit der Riemannschen Zetafunktion überein.

Kehren wir uns nunmehr dem endlich-Abelschen Fall zu. Mit $G$ bezeichnen wir eine endliche Abelsche Gruppe. Die $A_1, \dots, A_n$ in (1) sollen beliebige Untergruppen von $G$ bedeuten. Natürlich ist (1) so zu verstehen, daß im Fall $n=0$

$$\varrho(z; A_1, \dots, A_n) = 1 \qquad (n=0) \tag{4}$$

gilt. Man beachte, daß alle $A_1, \dots, A_n$ Untergruppen von $A_1 \dots A_n$ sind, weshalb man sich stets auf den Fall $G = A_1 \dots A_n$ beschränken dürfte, was wir doch der größeren Allgemeinheit halber nicht tun wollen.

Die Definition (1) ist jetzt für alle komplexen $z$ sinnvoll, und zwar ist $\varrho(z)$ eine ganze transzendente Funktion. Schon aus diesem Grunde empfiehlt sich, durchweg $\varrho(z)$ statt $\zeta(z) = \varrho(z)^{-1}$ zu untersuchen. Hierdurch werden sich auch unsere Formeln einfacher gestalten.

Man bemerke

$$\varrho(0; A_1, \dots, A_n) = 0 \qquad (n \geqq 1). \tag{5}$$

Wir schreiben (1) auch in einer mehr expliziten Form hin:

$$\varrho(z) = \varrho(z; A_1, \dots, A_n) = \sum_{k=0}^{n} \sum_{1 \leqq i_1 \leqq \dots \leqq i_k \leqq n} (-1)^k (A_{i_1} \dots A_{i_k})^{-z}, \tag{6}$$

wobei wir berücksichtigt haben, daß das Produkt $A_{i_1} \dots A_{i_k}$ eben die durch die Elemente der Faktoren erzeugte Untergruppe von $G$ ist. (Auch für die nichtabelschen $G$ gilt (6) mit der Änderung, daß man statt $A_{i_1} \dots A_{i_k}$ mit gewohnter Bezeichnung $\{A_{i_1}, \dots, A_{i_k}\}$ schreibt.)

Zum Beispiel lautet (6) für $n = 1, 2, 3$ so:

$$\varrho(z; A) = 1 - \frac{1}{(A)^z}, \tag{7}$$

$$\varrho(z; A, B) = 1 - \frac{1}{(A)^z} - \frac{1}{(B)^z} + \frac{1}{(AB)^z}, \tag{8}$$

$$\varrho(z; A, B, C) = 1 - \frac{1}{(A)^z} - \frac{1}{(B)^z} - \frac{1}{(C)^z} + \frac{1}{(AB)^z} + \frac{1}{(AC)^z} + \frac{1}{(BC)^z} - \frac{1}{(ABC)^z}. \tag{9}$$

An diese Beispiele wollen wir einige Bemerkungen knüpfen. Den sehr einfachen Fall (7) außer Acht gelassen, bemerken wir, daß auch Fall (8) kein nennenswertes Problem bietet. Bekanntlich gilt nämlich

$$(AB) = \frac{(A)(B)}{(A \cap B)}. \tag{10}$$

Hieraus ersieht man leicht, daß die Tripel $(A), (B), (AB)$ mit den sämtlichen Tripeln $ad, bd, abd$ zusammenfallen, wo die Parameter $a, b, d$ voneinander unabhängig alle natürlichen Zahlen durchlaufen. Hiernach stimmen die $\varrho(z; A, B)$ in ihrer Gesamtheit mit den Funktionen

$$1 - (ad)^{-z} - (bd)^{-z} + (abd)^{-z} \qquad (a, b, d = 1, 2, \dots)$$

überein.

Dagegen ist es um (9) nicht mehr so einfach. Vor allem gilt nämlich für $(ABC)$ eine ähnliche Formel wie (10) nicht. So z. B. ist im allgemeinen

$$(ABC) \neq \frac{(A)(B)(C)(A \cap B \cap C)}{(A \cap B)(A \cap C)(B \cap C)},$$

wovon man sich leicht überzeugen kann. Überhaupt scheint uns, daß die in (9) vorkommenden sieben Gruppenordnungen $(A), (B), (C), (AB), (AC), (BC), (ABC)$ ein nicht leicht übersehbares System bilden, weshalb von der Funktion (9) kein triviales Verhalten zu erwarten ist. Noch mehr bezieht sich das auf (1) im allgemeinen.

Fassen wir jetzt nur die reellen $z(\geqq 1)$ ins Auge. Von (8) liest man sofort ab, daß

$$0 \leqq \varrho(z; A, B) < 1 \qquad (z \geqq 1)$$

gilt. Ferner sieht man, daß „=" nur dann gilt, wenn mindestens das eine von $(A)$ und $(B)$ gleich 1 (und $z$ beliebig) ist. Mit einiger Mühe bekommt man, daß auch

$$0 \leqq \varrho(z; A, B, C) < 1 \qquad (z \geqq 1)$$

zutrifft, daß aber „=" außer den trivialen Fällen „$(A)$ oder $(B)$ oder $(C)$ gleich 1 und $z$ beliebig" auch noch im Fall gilt, wo $z = 1$ ist und $A, B, C$ die drei echten Untergruppen der nichtzyklischen Gruppe vierter Ordnung sind. In diesem Zusammenhang bemerken wir gleich schon hier, daß im allgemeinen die Untersuchung der Nullstellen von $\varrho(z; A_1, \ldots, A_n)$ für $z = 1, 2, \ldots$ sich für ein sehr interessantes Problem erweisen wird (s. unten).

Man würde allgemein zunächst nur fragen, ob es absolute Konstanten $c, C(>0)$ gibt, so daß stets

$$-c < \varrho(z; A_1, \ldots, A_n) < C \qquad (z \geqq 1) \tag{11}$$

gilt. Die Antwort wird verneinend, aber insbesondere für die ganzen $z$ gewinnen wir die überraschenden Ungleichungen

$$0 \leqq \varrho(z; A_1, \ldots, A_n) < 1 \qquad (n \geqq 1;\ z = 1, 2, \ldots), \tag{12}$$

die wir als ein sehr bemerkenswertestes Resultat ansehen und den *Trägheitssatz* der Theorie der endlichen Abelschen Gruppen nennen. Um diesen zu würdigen bemerken wir, daß es sich in ihm um einen in dreierlei Hinsicht scharfen Satz handelt. *Erstens* gilt nämlich (12) für die endlichen nichtabelschen Gruppen nicht mehr, sogar ist für diese die linke Seite von (12) mit jedem $c$ falsch, weshalb (12) eine echt Abelsche Tatsache ist. *Zweitens* ist (12) auch in dem Sinne scharf, daß in ihm sehr oft, und zwar in keineswegs trivialen Fällen „=" gilt. *Drittens* ist (12) auch in der Beziehung scharf, daß es für die reellen aber nichtganzen $z(>1)$ nicht mehr allgemein gilt, was wir für sehr auffallend finden; es wird sich sogar zeigen, daß dann beide Seiten von (11) falsch sind, wie man auch $c, C$ wählt.

Eine interessante Folgerung aus obigem Trägheitssatz wird sein, daß zwischen unseren Zetafunktionen und dem Satz von HAJÓS [2][2] über die endlichen Abelschen Gruppen ein enger Zusammenhang besteht. Und zwar läßt sich dieser tiefliegende Satz in einer äquivalenten Form als die notwendige und hinreichende Bedingung aussprechen, damit gewisse Zetafunktionen $\zeta(z)$ an der Stelle $z=1$ einen Pol, d. h. die entsprechenden reziproken Funktionen $\varrho(z)=\zeta(z)^{-1}$ eine Nullstelle haben.

Das gesagte beleuchtet die Bedeutung unserer Zetafunktionen genügend, weshalb wir meinen, daß durch sie in der Algebra im allgemeinen, insbesondere aber jedenfalls in der Theorie der endlichen Abelschen Gruppen ein wichtiges Forschungsgebiet eröffnet wird.

Im § 2 stellen wir unsere Resultate bezüglich der gruppentheoretischen Zetafunktionen im endlich-Abelschen Fall zusammen (auf den Zusammenhang mit dem Satz von HAJÓS gehen wir aber erst in einer anderen Arbeit [3] ein). Im § 3 erfolgen die Beweise. Im § 4 machen wir einige Schlußbemerkungen auch über den allgemeinen Fall.

## § 2. Der endlich-Abelsche Fall

Mit ausnahme vom § 4 soll eine endliche Abelsche Gruppe $G$ mit dem Einselement $\varepsilon$ vorgelegt sein, ferner sollen $A_1, \dots, A_n$ durchweg Untergruppen von $G$ bezeichnen, die im allgemeinen nicht verschieden zu sein brauchen. Alle unseren nachfolgenden Sätze formulieren wir bequemlichkeitshalber bezüglich $\varrho(z)=\zeta^{-1}(z)$ statt $\zeta(z)$. Die Sätze 1 bis 7 werden fast trivial sein, dagegen drücken die Sätze 8 bis 12 tiefliegende Sachverhalte aus.

SATZ 1. (Vertauschungsregel.) *Es ist $\varrho(z; A_1, \dots, A_n)$ von der Reihenfolge der $A_n$ unabhängig.*

SATZ 2. (Kürzungsregel.) *Enthält $A_1$ ein $A_i$ $(i=2, \dots, n)$, so gilt*

$$\varrho(z; A_1, \dots, A_n)=\varrho(z; A_2, \dots, A_n) \qquad (n \geqq 2). \tag{13}$$

KOROLLAR. *Behält man aus den Untergruppen $A_1, \dots, A_n$ zunächst nur die verschiedenen und dann aus diesen nur die minimalen (d. h. die keine der übrigen enthalten), so gilt für die übriggebliebenen Untergruppen $B_1, \dots, B_r$:*

$$\varrho(z; A_1, \dots, A_n)=\varrho(z; B_1, \dots, B_r).$$

*Insbesondere wenn ein $A_i$ gleich $\varepsilon$ ist, so gilt*

$$\varrho(z; A_1, \dots, A_n)=0. \tag{14}$$

(Es ist klar, daß umgekehrt $\varrho(z)=\varrho(z; A_1, \dots, A_n)$ $(n \geqq 1)$ nichtkonstant ist, wenn alle $A_i$ von $\varepsilon$ verschieden sind, da dann $\varrho(+\infty)=1$ ist, dagegen nach (5) $\varrho(0)=0$ gilt.)

[2] Mit [ ] verweisen wir auf das Literaturverzeichnis am Ende unserer Arbeit.

SATZ 3. (Erste additive Eigenschaft.) *Für jede Untergruppe $B$ von $A_1$ gilt*

$$\varrho(z; A_1, \ldots, A_n) = \varrho(z; B, A_2, \ldots, A_n) + \\ + (B)^{-z} \varrho(z; A_1/B, A_2B/B, \ldots, A_nB/B) \qquad (n \geqq 1; B \subseteqq A_1), \tag{15}$$

*wobei „/" die Faktorgruppe bezeichnet,*

KOROLLAR. *Sind die $A_i$ von $\varepsilon$ verschieden, so gilt*

$$\varrho(z; A_1, \ldots, A_n) = \sum_{B_1, \ldots, B_n} (B)^{-z} \varrho(z; B_1'B/B, \ldots, B_n'B/B) \qquad (n \geqq 1), \tag{15'}$$

*wobei $B_i$ die von $A_i$ verschiedenen Glieder einer beliebigen Kompositionsreihe von $A_i$ durchläuft, $B_i'$ das $B_i$ enthaltende benachbarte Glied dieser Kompositionsreihe bezeichnet $(i = 1, \ldots, n)$ und $B$ für $B_1 \ldots B_n$ steht.* (Es genügt diejenigen Glieder der rechten Seite von (15′) beizubehalten, in denen $B$ kein $B_i'$ enthält, da die übrigen Glieder wegen (14) verschwinden. In den verbleibenden Gliedern sind alle Gruppen $B_i'B/B$ offenbar von Primzahlordnung.)

SATZ 4. (Zweite additive Eigenschaft.) *Es gilt*

$$\varrho(z; A_2, \ldots, A_n) = \varrho(z; A_1, \ldots, A_n) + \\ + (A_1)^{-z} \varrho(z; A_1A_2/A_1, \ldots, A_1A_n/A_1) \qquad (n \geqq 1). \tag{16}$$

SATZ 5. (Dritte additive Eigenschaft.) *Es gilt*

$$\sum_{\mathfrak{M} \subseteqq \mathfrak{N}} (A_{\mathfrak{M}})^{-z} \varrho(z; A_x A_{\mathfrak{M}}/A_{\mathfrak{M}}, A_y A_{\mathfrak{M}}/A_{\mathfrak{M}}, \ldots) = 1, \tag{17}$$

*wobei $x, y, \ldots$ jedesmal die sämtlichen Elemente der Menge $\mathfrak{N} - \mathfrak{M}$ bezeichnen.* (Man bemerke, daß die linke Seite von (17) insbesondere die den Fällen $\mathfrak{M} = \mathbf{0}, \mathfrak{N}$ entsprechenden Glieder $\varrho(z; A_1, \ldots, A_n)\,(A_1 \ldots A_n)^{-z}$ enthält.)

Zur Beleuchtung von (17) schreiben wir den Fall $n = 3$ explizit hin:

$$\varrho(z; A, B, C) + (A)^{-z}\varrho(z; AB/A, AC/A) + (B)^{-z}\varrho(z; AB/B, BC/B) + \\ + (C)^{-z}\varrho(z; AC/C, BC/C) + (AB)^{-z}\varrho(z; ABC/AB) + (AC)^{-z}\varrho(z; ABC/AC) + \\ + (BC)^{-z}\varrho(z; ABC/BC) + (ABC)^{-z} = 1.$$

SATZ 6. (Multiplikative Eigenschaft.) *Haben die Produkte $A_1 \ldots A_i$, $A_{i+1} \ldots A_n$ nur das Element $\varepsilon$ gemein, so gilt*

$$\varrho(z; A_1, \ldots, A_n) = \varrho(z; A_1, \ldots, A_i)\, \varrho(z; A_{i+1}, \ldots, A_n). \tag{18}$$

SATZ 7. (Additiv-multiplikative Eigenschaft.) *Es gilt*

$$\varrho(z; A_1, \ldots, A_n) = \varrho(z; A_2, \ldots, A_n) + \varrho(z; A_1, A_3, \ldots, A_n) - \\ - \varrho(z; A_1A_2, A_3, \ldots, A_n) \qquad (n \geqq 2). \tag{19}$$

Um den Satz 8 formulieren zu können, definieren wir eine Funktion $\psi(z, G)$ folgenderweise. Ist $G = P$ eine $p$-Gruppe und $t$ die Anzahl ihrer Invarianten, so sei

$$\psi(z, P) = \prod_{l=0}^{t-1} (1 - p^{l-z}) = (1 - p^{-z})(1 - p^{1-z}) \ldots (1 - p^{t-1-z}) \qquad (t \geqq 1). \tag{20}$$

Ist ferner $G$ das direkte Produkt von $P_1, \ldots, P_s$, wobei $P_k$ eine $p_k$-Gruppe ist und $p_1, \ldots, p_s$ verschiedene Primzahlen sind, so sei

$$\psi(z, G) = \prod_{k=1}^{s} \psi(z, P_k). \tag{21}$$

(Im Fall $G = \varepsilon$ ist natürlich $\psi(z, G) = 1$ zu verstehen. Im Fall $G \neq \varepsilon$ ist $\psi(z, G)$ nichtkonstant.)

SATZ 8. (Hauptsatz der Zetafunktionen für endliche Abelsche Gruppen.) *Sind $A_1, \ldots, A_n$ Untergruppen der endlichen Abelschen Gruppe $G$, so gilt*

$$\varrho(z; A_1, \ldots, A_n) = \sum_{A_1, \ldots, A_n \not\subseteq H \subseteq G} (H)^{-z} \psi(z, G/H) \qquad (n \geqq 1), \tag{22}$$

*wobei man über alle Untergruppen $H$ von $G$ zu summieren hat, die kein $A_i$ enthalten (somit von $G$ verschieden sind).*

BEMERKUNG 1. Aus diesem Satz folgt, daß die rechte Seite von (22) für alle $G$ gleich ist, die das Produkt $A_1 \ldots A_n$ enthalten, was man leicht auch direkt beweisen könnte. Hat man es mit nur einem $\varrho(z; A_1, \ldots, A_n)$ zu tun, so wendet man (22) am bequemsten mit $G = A_1 \ldots A_n$ an. Auch die allgemeinere Annahme $G \supseteq A_1 \ldots A_n$ kann in (22) von Vorteil sein, wenn man nämlich gleichzeitig mehrere $\varrho(z; A_1, \ldots, A_n)$ betrachtet, in denen die $A_i$ irgendwelche Untergruppen einer gegebenen endlichen Abelschen Gruppe $G$ sind.

BEMERKUNG 2. Ist ein $A_i$ gleich $\varepsilon$, so ist die Summe in (22) leer, weshalb dann (22) eben in (14) übergeht. Sind dagegen alle $A_i$ von $\varepsilon$ verschieden, so ist in (22) der Fall $H = \varepsilon$ gewiß zuzulassen. Hiernach folgt aus (22):

$$\varrho(z; A_1, \ldots, A_n) = \psi(G) + \sum_{A_1, \ldots, A_n \not\subseteq H \subset G, \neq \varepsilon} (H)^{-z} \psi(z, G/H) \tag{23}$$
$$(n \geqq 1; A_1, \ldots, A_n \neq \varepsilon).$$

BEMERKUNG 3. Sind insbesondere alle $A_i$ von $p$-ter Ordnung ($p$ Primzahl), so folgt aus (20), (23) (mit Anwendung auf $G = A_1 \ldots A_n$) offenbar

$$\varrho(z; A_1, \ldots, A_n) = \prod_{l=0}^{t-1} (1 - p^{l-z}) + \sum_{k=1}^{t-1} \nu_k p^{-kz} \prod_{l=0}^{t-k-1} (1 - p^{l-z}) \tag{24}$$
$$(n \geqq 1; (A_1) = \cdots = (A_n) = p; p^t = (A_1 \ldots A_n)),$$

wobei $\nu_k$ die Anzahl derjenigen Untergruppen $p^k$-ter Ordnung von $A_1 \ldots A_n$ bezeichnet, die kein $A_i$ enthalten. Das erste Produkt in (24) verschwindet dann und nur dann, wenn $z$ eine der ganze Zahlen $0, \ldots, t-1$ ist; in diesen Fällen verschwindet auch das zweite Produkt stets, wenn außerdem $t-k-1 \geqq z$, d. h. $k \leqq t-z-1$ gilt, weshalb aus (24) folgt

$$\varrho(z; A_1, \ldots, A_n) = \sum_{k=t-z}^{t-1} \nu_k p^{-kz} \prod_{l=0}^{t-k-1} (1 - p^{l-z}) \tag{25}$$
$$(n \geqq 1; (A_1), \ldots, (A_n) | p; p^t = (A_1 \ldots A_n); z = 0, \ldots, t-1),$$

wobei auch der Fall zugelassen werden durfte, daß einige der $A_i$ gleich $\varepsilon$ sind, da dann $\nu_1 = \cdots = \nu_{t-1} = 0$ ist, somit beide Seiten von (25) verschwinden. Z. B. für $z = 1, 2$ gelten also

$$(26)\qquad \varrho(1; A_1, \ldots, A_n) = \frac{(p-1)\nu_{t-1}}{p^t} \qquad (t \geqq 2)$$

$$(27)\qquad \varrho(2; A_1, \ldots, A_n) = \frac{(p^2-1)\nu_{t-1} + (p^2-1)(p^2-p)\nu_{t-2}}{p^{2t}} \qquad (t \geqq 3)$$

$$((A_1), \ldots, (A_n) | p; p^t = (A_1 \ldots A_n)).$$

Bemerkung 4. Wir nennen Satz 8 deshalb einen Hauptsatz, weil er für die Berechnung von $\varrho(z)$ eine der Definition (1) gegenüber wesentlich neue Möglichkeit bietet und einen Ausgangspunkt für weitere Schlüße bildet, worüber in den folgenden Sätzen 9 bis 12 die Rede sein wird.

Satz 9. (Trägheitssatz.) *Es gilt*

$$(28)\qquad 0 \leqq \varrho(z; A_1, \ldots, A_n) < 1 \qquad (n \geqq 1; z = 1, 2, \ldots).$$

Zusatz. *Die rechte Seite von* (28) *gilt sogar mit*

„$\leqq 1 - (\min((A_1), \ldots, (A_n)))^{-z}$“ *statt* „$< 1$“.

Bemerkung 5. Wir nennen Satz 9 einen Trägheitssatz, weil nach ihm die $\varrho(z)$ in (28) nur sehr kleiner Schwankungen fähig sind. In der Einleitung haben wir schon gesagt, daß Satz 9 in dreierlei Hinsicht „scharf“ ist. Das wollen wir hier mit Beispielen bestätigen. *Erstens* bedeute jetzt $G$ die Diedergruppe von der Ordnung $2p$ ($p$ ungerade Primzahl). Diese hat $p$ Untergruppen 2-ter Ordnung, die wir mit $A_1, \ldots, A_p$ bezeichnen. Wir wollen $\varrho(z; A_1, \ldots, A_p)$ berechnen. Dies besteht nach (1) aus gleichvielen Gliedern vom Vorzeichen „$+$“ bzw. „$-$“. Eins davon ist 1, weitere $p$ Glieder sind gleich $-2^{-z}$, alle übrigen sind gleich $\pm(2p)^{-z}$, weshalb offenbar

$$\varrho(z; A_1, \ldots, A_p) = 1 - \frac{p}{2^z} + \frac{p-1}{(2p)^z}$$

gilt. Man sieht, daß die rechte Seite für kein $z = 1, 2, \ldots$ (sogar für kein $z \geqq 1$) von unten beschränkt ist, weshalb (28) für die endlichen *nichtabelschen* Gruppen falsch ist. *Zweitens* bezeichne $G$ eine elementare Abelsche Gruppe von der Ordnung $p^t$ ($p$ Primzahl, $t \geqq 2$). Sind dann $A_1, \ldots, A_n$ die sämtlichen Untergruppen $p$-ter Ordnung von $G$, so sind in (25) alle $\nu_k$ gleich 0, weshalb jetzt in (28) für $z = 1, \ldots, t-1$ das Zeichen „$=$“ gilt. Man könnte mit Leichtigkeit ähnliche aber weniger einfache Beispiele in großer Zahl konstruieren (s. auch Bemerkung 6 und Satz 10). *Drittens* betrachten wir wieder das vorige $\varrho(z; A_1, \ldots, A_n)$, jetzt aber für die reellen $z (\geqq 1)$. Nach (24) gilt

$$\varrho(z; A_1, \ldots, A_n) = \prod_{l=0}^{t-1} (1 - p^{l-z}).$$

Man sieht, daß dies für die reellen $z$ im Innern der Intervalle $(1,2), (2,3), \ldots, (t-2, t-1)$ vom abwechselnden Vorzeichen ist, daß also die linke Seite von (28) für die nichtganzen $z(>1)$ im allgemeinen falsch ist; da $t$ beliebig groß sein kann, so sieht man sogar, daß — wie auch $c, C$ gewählt werden — keine Seite von (11) allgemein gilt.

Bemerkung 6. Wir wollen genauer zusehen, was für Möglichkeiten vorliegen, damit eine positive ganze Zahl $z$ eine Nullstelle von einem $\varrho(z)$, d. h. ein Pol der zugehörigen Zetafunktion $\zeta(z) = \varrho(z)^{-1}$ ist. (Diese Frage wird nicht nur an sich, sondern — wie in der Einleitung erwähnt wurde — auch im Zusammenhang mit dem Satz von Hajós interessant.) Man liest von (20) ab, daß $\psi(z, P)$ für $z = 0, 1, \ldots$ nichtnegativ, und zwar für $z = 0, \ldots, t-1$ gleich 0, für $z = t, t+1, \ldots$ positiv ist. Ähnliches gilt nach (21) auch über $\psi(z, G)$, wenn wir $t$ folgenderweise definieren:

$$t = t(G) = \max(t_1, \ldots, t_s), \tag{29}$$

wobei $t_i$ die Anzahl der Invarianten des Faktors $P_i$ in der Darstellung von $G$ als ein direktes Produkt $P_1 \ldots P_s$ von $p_i$-Gruppen $P_i$ bedeutet ($p_1, \ldots, p_s$ verschiedene Primzahlen). Man bemerke auch, daß stets $t(G/H) \leqq t(G)$ gilt. Nun gilt nach dem Gesagten, daß der Summand in (22) für $z = 0, 1, \ldots$ nichtnegativ, und zwar für $z = 0, \ldots, t(G/H)-1$ gleich 0, für $z = t(G/H), t(G/H)+1, \ldots$ positiv ist. So kommen wir zum folgenden:

Satz 10. *Um die positiven ganzzahligen Nullstellen von $\varrho(z; A_1, \ldots, A_n)$ zu bestimmen, wobei $A_1, \ldots, A_n$ $(n \geqq 1)$ von 1 verschiedene Untergruppen einer endlichen Abelschen Gruppe $G$ sind, verfahre man so: Man nehme alle Untergruppen $H$ von $G$, die kein $A_i$ enthalten, bilde für sie*

$$t_0 = \min_H t(G/H), \tag{30}$$

*wobei $t(G/H)$ nach* (29) *zu verstehen ist. Dann verschwindet $\varrho(z; A_1, \ldots, A_n)$ für $z = 1, \ldots, t_0 - 1$ und ist für $z = t_0, t_0+1, \ldots$ positiv.*

Korollar. *Die sämtlichen positiven ganzen Nullstellen eines nichtkonstanten $\varrho(z; A_1, \ldots, A_n)$ bilden eine lückenlose Folge $1, \ldots, t_0 - 1$, wobei für $t_0$* (30) *gilt.* (Im Fall $t_0 = 1$, und nur in diesem, gibt es keine solchen Nullstellen.)

Bemerkung 7. Wir formulieren eine leichte Folgerung vom Satz 10 besonders: *Für beliebige Untergruppen $A_1, \ldots, A_n$ von $G$ gilt $\varrho(1; A_1, \ldots, A_n) = 0$ dann und nur dann, wenn jede Untergruppe $H$ von $G$ mit zyklischer Faktorgruppe $G/H$ mindestens ein $A_i$ enthält.* (Hierbei darf unter den $A_i$ auch $\varepsilon$ vorkommen.)

Satz 11. (Monotone Eigenschaft.) *Stets gilt*

$$\varrho(z; A_1, \ldots, A_n) \leqq \varrho(z; A_1, \ldots, A_{n-1}) \qquad (n \geqq 1; z = 1, 2, \ldots). \tag{31}$$

Die vorigen bilden die Grundlagen einer elementaren Theorie der betrachteten Zetafunktionen. Wir haben nicht untersucht, wie weit für sie durch Einsetzung von funktionentheoretischen Hilfsmitteln die Begründung einer höheren Theorie möglich ist. Leicht scheint das wegen der sich in der freien Wahl des Untergruppensystems $A_1, \ldots, A_n$ offenbarenden, ungeheuer großen Allgemeinheit unserer Zetafunktionen nicht zu sein. Hier wollen wir nur eine ganz leichte funktionentheoretische Anwendung machen, die so entstehen wird, daß man den Differentialquotienten $\varrho'(0)$ sowohl nach (1), als auch nach (22) berechnet und die zwei Resultate einander gleichsetzt. So werden wir zum folgenden Satz kommen:

Satz 12. *Man definiere zu (1) ähnlich das Produkt*

$$\pi(A_1, \ldots, A_n) = \prod_{\mathfrak{M} \subset \mathfrak{N}} (A_{\mathfrak{M}})^{(-1)^{(\mathfrak{M})}} \qquad (n \geqq 1), \tag{32}$$

*wobei $A_1, \ldots, A_n$ Untergruppen einer endlichen Abelschen Gruppe $G$ sind. Bezeichnen $p_1, \ldots, p_s$ die verschiedenen Primteiler in der Ordnung von $G$, so gilt*

$$\pi(A_1, \ldots, A_n) = \prod_{k=1}^{s} p_k^{-\sum'_{H_k} \prod_{l=1}^{t(G/H_k)-1} (1-p_k^l)}, \tag{33}$$

*wobei $H_k$ diejenigen Untergruppen von $G$ durchläuft, die kein $A_i$ enthalten und für die die Faktorgruppe $G/H_k$ eine $p_k$-Gruppe ist, ferner $t(G/H_k)$ die Anzahl der Invarianten von $G/H_k$ bezeichnet.*

Man bemerke, daß (33) beliebig groß sein und auch beliebig nahe an 0 liegen kann. Ist nämlich $G$ eine elementare Abelsche $p$-Gruppe von der Ordnung $p^t$ und sind $A_1, \ldots, A_n$ die sämtlichen Untergruppen $p$-ter Ordnung von $G$, so geht (33) in

$$\pi(A_1, \ldots, A_n) = p^{-(1-p)(1-p^2)\ldots(1-p^{t-1})}$$

über, woraus die Behauptung folgt.

## § 3. Beweis

Der Satz 1 ist eine Folgerung aus der Definition (1).

Für das nächtsfolgende bezeichnen wir mit $\mathfrak{N}_1$ die Menge der Zahlen $2, \ldots, n$ $(n \geqq 1)$; insbesondere für $n = 1$ ist also $\mathfrak{N}_1 = 0$. Dann gilt nach (1) offenbar

$$\varrho(z; A_1, \ldots, A_n) = \sum_{\mathfrak{M} \subseteq \mathfrak{N}_1} (-1)^{(\mathfrak{M})} (A_{\mathfrak{M}})^{-z} - \sum_{\mathfrak{M} \subseteq \mathfrak{N}_1} (-1)^{(\mathfrak{M})} (A_1 A_{\mathfrak{M}})^{-z} \quad (n \geqq 1). \tag{34}$$

Um nunmehr Satz 2 zu beweisen dürfen wir wegen Satz 1 annehmen, daß z. B. $A_1 \supseteq A_2$ gilt. Wir fassen die sämtlichen $\mathfrak{M} (\subseteq \mathfrak{N}_1)$ in Paare $\mathfrak{M}_1, \mathfrak{M}_2$,

so daß jedesmal $\mathfrak{M}_1, \mathfrak{M}_2$ voneinander nur im Element 2 unterscheiden. Dann gilt $(\mathfrak{M}_1)-(\mathfrak{M}_2)=\pm 1$, $A_1 A_{\mathfrak{M}_1}=A_1 A_{\mathfrak{M}_2}$, folglich verschwindet jetzt die zweite Summe auf der rechten Seite von (34). Hiernach geht (34) in (13) über, womit Satz 2 bewiesen ist.

Die erste Hälfte des Korollars folgt unmittelbar aus Satz 2. Da für $A=\varepsilon$ offenbar $\varrho(z; A)=0$ gilt, so ist auch (14) richtig.

Zum Beweis von Satz 3 wenden wir (34) mit $B$ statt $A_1$ an. Nach Subtraktion beider Gleichungen ergibt sich

$$\varrho(z; A_1, \ldots, A_n)-\varrho(z; B, A_2, \ldots, A_n)=\sum_{\mathfrak{M}\subset\mathfrak{N}_1}(-1)^{(\mathfrak{M})}((BA_{\mathfrak{M}})^{-z}-(A_1 A_{\mathfrak{M}})^{-z}).$$

Wenn nun $B\subseteqq A_1$ gilt, so können wir die rechte Seite in

$$(B)^{-z}\sum_{\mathfrak{M}\subset\mathfrak{N}_1}(-1)^{(\mathfrak{M})}((BA_{\mathfrak{M}}/B)^{-z}-(A_1 A_{\mathfrak{M}}\, B)^{-z})$$

umformen. Dies ist wegen $A_1=BA_1$ offenbar gleich

$$(B)^{-z}\sum_{\mathfrak{M}\subseteq\mathfrak{N}}(-1)^{(\mathfrak{M})}(BA_{\mathfrak{M}}/B)^{-z},$$

d. h. nach (1) gleich

$$(B)^{-z}\varrho(z; BA_1/B, \ldots, BA_n/B).$$

Nach Berücksichtigung von $BA_1=A_1$ folgt hieraus Satz 3.

Das Korollar ließe sich aus Satz 3 durch Induktion gewinnen, bequemer beweisen wir es aber direkt. Die rechte Seite von (15′) läßt sich nach (1) so schreiben:

$$\sum_{B_1, \ldots, B_n}\sum_{\mathfrak{M}\subseteq\mathfrak{N}}(B)^{-z}(-1)^{(\mathfrak{M})}(B'_{\mathfrak{M}}B/B)^{-z},$$

wobei $B'_{\mathfrak{M}}$ das Produkt der $B'_i$ $(i\in\mathfrak{M})$ bezeichnet. Hierfür schreibt sich:

$$\sum_{B_1,\ldots,B_n}\sum_{\mathfrak{M}\subseteq\mathfrak{N}}(-1)^{(\mathfrak{M})}(B'_{\mathfrak{M}}B)^{-z}.$$

Es ist

$$B'_{\mathfrak{M}}B=B'_{\mathfrak{M}}B_1, \ldots, B_n=B'_{\mathfrak{M}}B_{\mathfrak{N}-\mathfrak{M}},$$

wobei $B_{\mathfrak{K}}$ für $\mathfrak{K}\subset\mathfrak{M}$ das Produkt der $B_i$ $(i\in\mathfrak{K})$ bezeichnet. Hiernach geht die vorige Summe in

$$\sum_{B_1, \ldots, B_n}\sum_{\mathfrak{M}\subseteq\mathfrak{N}}(-1)^{(\mathfrak{M})}(B'_{\mathfrak{M}}B_{\mathfrak{N}-\mathfrak{M}})^{-z}$$

über. Jedes $B'_{\mathfrak{M}}B_{\mathfrak{N}-\mathfrak{M}}$ entsteht so aus $B_1, \ldots, B_n$, daß man die $B_i$ $(i\in\mathfrak{M})$ durch $B'_i$ ersetzt. Hieraus ist leicht zu sehen, daß die vorige Summe nach Streichung entgegengesetzt gleicher Glieder in

$$\sum_{\mathfrak{M}\subseteq\mathfrak{N}}(-1)^{(\mathfrak{M})}(A_{\mathfrak{M}})^{-z}$$

übergeht. Das besagt nach (1) die Richtigkeit von (15′), womit wir das Korollar vom Satz 3 bewiesen haben.

Satz 4 ist im Fall $n=1$ trivial. Der Fall $n\geqq 2$ entsteht aus dem Fall $n=2$ vom Satz 3 so, daß man diesen für $A_1, A_2, A_1$ statt $A_1, B_1$ anwendet und Satz 2 berücksichtigt.

Um Satz 5 zu beweisen drücken wir die linke Seite von (17) nach (1) so aus:

$$\sum_{\mathfrak{M}\subset\mathfrak{N}}(A_{\mathfrak{M}})^{-z}\sum_{\mathfrak{M}_1\subset\mathfrak{N}-\mathfrak{M}}(-1)^{(\mathfrak{M}_1)}(A_{\mathfrak{M}_1}A_{\mathfrak{M}}/A_{\mathfrak{M}})^{-z}.$$

Hierfür schreibt sich

$$\sum_{\mathfrak{M}\subset\mathfrak{N}}(A_{\mathfrak{M}})^{-z}\sum_{\mathfrak{M}_1\subseteq\mathfrak{N}-\mathfrak{M}}(-1)^{(\mathfrak{M}_1)}(A_{\mathfrak{M}_1}A_{\mathfrak{M}})^{-z}.$$

Nach Umordnung der Glieder entsteht hieraus (wegen $A_{\mathfrak{M}_1}A_{\mathfrak{M}}=A_{\mathfrak{M}_1\cup\mathfrak{M}}$)

$$\sum_{\mathfrak{A}\subseteq\mathfrak{N}}\sum_{\mathfrak{B}\subseteq\mathfrak{A}}(-1)^{(\mathfrak{B})}(A_{\mathfrak{A}})^{-z}.$$

Die innere Summe ist 1, wenn $\mathfrak{A}=0$ ist, und verschwindet für $\mathfrak{A}\neq 0$. Damit haben wir Satz 5 bewiesen.

Zum Beweis von Satz 6 zerlegen wir $\mathfrak{N}$ in zwei komplementäre Untermengen $\mathfrak{N}', \mathfrak{N}''$, so daß diese aus $1,\ldots,i$ bzw. $i+1,\ldots,n$ bestehen. Nach (1) gilt dann

$$\varrho(z; A_1,\ldots,A_n)=\sum_{\mathfrak{M}_1\subseteq\mathfrak{N}'}\ \sum_{\mathfrak{M}_2\subseteq\mathfrak{N}''}(-1)^{(\mathfrak{M}_1)+(\mathfrak{M}_2)}(A_{\mathfrak{M}_1\cup\mathfrak{M}_2})^{-z}. \tag{35}$$

Es ist $A_{\mathfrak{M}_1\cup\mathfrak{M}_2}=A_{\mathfrak{M}_1}A_{\mathfrak{M}_2}$. Andererseits gilt $A_{\mathfrak{M}_1}\subseteq A_1\ldots A_i$, $A_{\mathfrak{M}_2}\subseteq A_1\ldots A_n$, somit folgt aus der Annahme: $(A_{\mathfrak{M}_1\cup\mathfrak{M}_2})=(A_{\mathfrak{M}_1})(A_{\mathfrak{M}_2})$. Hiernach geht (35) in (18) über, womit Satz 6 bewiesen ist.

Um Satz 7 zu beweisen bezeichnen wir mit $\mathfrak{N}_2$ die Menge der Zahlen $3,\ldots,n$ $(n\geqq 2)$. Nach (1) gilt dann

$$\varrho(z; A_1,\ldots,A_n)=\sum_{\mathfrak{M}\subseteq\mathfrak{N}}(-1)^{(\mathfrak{M})}((A_{\mathfrak{M}})^{-z}-(A_1A_{\mathfrak{M}})^{-z}-(A_2A_{\mathfrak{M}})^{-z}+(A_1A_2A_{\mathfrak{M}})^{-z}).$$

Ähnlich gelten

$$\varrho(z; A_1, A_2, A_3,\ldots,A_n)=\sum_{\mathfrak{M}\subset\mathfrak{N}_2}(-1)^{(\mathfrak{M})}((A_{\mathfrak{M}})^{-z}-(A_1A_2A_{\mathfrak{M}})^{-z}),$$

$$\varrho(z; A_i, A_3,\ldots,A_n)=\sum_{\mathfrak{M}\subset\mathfrak{N}_2}(-1)^{(\mathfrak{M})}((A_{\mathfrak{M}})^{-z}-(A_iA_{\mathfrak{M}})^{-z})\qquad (i=1,2).$$

Aus diesen Gleichungen folgt Satz 7.

Wir wollen Satz 8 beweisen. Da die rechte Seite von (22) und wegen des Korollars vom Satz 2 auch die linke Seite von der Multiplizität der $A_i$ unabhängig ist, so dürfen wir annehmen, daß alle $A_i$ verschieden sind.

Für jede Untergruppe $K$ von $G$ bezeichnen wir mit $N_k(K)$ die Anzahl derjenigen Mengen $\mathfrak{M}(\subseteq\mathfrak{N})$, für die

$$(\mathfrak{M})=k,\quad A_{\mathfrak{M}}=K\qquad (K\subseteq G;\ k=0,\ldots,n) \tag{36}$$

ist. Nach (1) gilt dann

$$\varrho(z)=\varrho(z; A_1,\ldots,A_n)=\sum_{k=0}^{n}\sum_{K\subseteq G}(-1)^k N_k(K)(K)^{-z}. \tag{37}$$

Mit $\mathfrak{K}$ bezeichnen wir die Menge derjenigen $A_i$, die in $K$ enthalten sind. Nimmt man aus $\mathfrak{K}$ auf jede Art $k$ Untergruppen heraus und bildet ihr Produkt, so entsteht jede Untergruppe $H$ von $K$ genau $N_k(H)$-mal. Hiernach gilt

$$\binom{(\mathfrak{K})}{k} = \sum_{H \subset K} N_k(H) \qquad (K \subseteq G;\ k = 0, \ldots, n), \tag{38}$$

wobei links der Binomialkoeffizient steht und die Summe über alle Untergruppen $H$ von $K$ zu erstrecken ist.

Wir betrachten mit DELSARTE [1] irgendzwei Funktionen $f(G), F(G)$, die für alle endlichen Abelschen Gruppen $G$ erklärt sind, so daß beide Funktionswerte in einen Modul gehören und nur vom Typus von $G$ abhängen (d. h. $f(G) = f(G'), F(G) = F(G')$ gilt, wenn $G, G'$ isomorph sird). Gilt für alle $G$

$$F(G) = \sum_{H \subseteq G} f(G),$$

wobei man über alle Untergruppen $H$ von $G$ zu summieren hat, so gilt nach dem wichtigen Satz von DELSARTE [1] folgende Verallgemeinerung der Möbiusschen Umkehrungsformel:

$$f(G) = \sum_{H \subseteq G} \mu(G/H)\, F(H),$$

wobei die Funktion $\mu$ von MÖBIUS—DELSARTE folgende Bedeutung hat: Ist $Q$ eine elementare endliche Abelsche $p$-Gruppe, d. h. das direkte Produkt von $t$ Gruppen $p$-ter Ordnung, so ist

$$\mu(Q) = (-1)^t p^{\binom{t}{2}} \qquad (t = 0, 1, \ldots).$$

Ist ferner $G$ das direkte Produkt von $Q_1, \ldots, Q_s$, wobei $Q_i$ eine elementare endliche Abelsche $p_i$-Gruppe ist und $p_1, \ldots, p_s$ verschiedene Primzahlen sind, so ist

$$\mu(G) = \mu(Q_1) \ldots \mu(Q_s).$$

Ist endlich $G$ eine sonstige endliche Abelsche Gruppe, d. h. eine solche, deren mindestens eine Invariante durch ein Primzahlquadrat teilbar ist, so ist $\mu(G) = 0$.

Hiernach folgt aus (38)

$$N_k(K) = \sum_{H \subset K} \mu(K/H) \binom{(\mathfrak{H})}{k} \qquad (K \subset G; k = 0, \ldots, n),$$

wobei $\mathfrak{H}$ die Menge der in $H$ enthaltenen $A_i$ bezeichnet. Nach Einsetzung in (37) entsteht

$$\varrho(z) = \sum_{k=0}^{n} \sum_{K \subset G} \sum_{H \subset K} (-1)^k \mu(K/H) \binom{(\mathfrak{H})}{k} (K)^{-z}. \tag{39}$$

Es ist

$$\sum_{k=0}^{n}\binom{(\mathfrak{H})}{k}$$

gleich 1, wenn $(\mathfrak{H})=0$ ist, d. h. wenn $H$ kein $A_i$ enthält. Im anderen Fall dagegen verschwindet diese Summe. Wird dies in (39) berücksichtigt, so entsteht nach leichter Umordnung der Glieder

$$\varrho(z)=\sum_{H}{}' \sum_{H\subseteq K\subset G} \mu(K/H)\,(K)^{-z},$$

wo man über die kein $A_i$ enthaltenden Untergruppen $H$ von $G$ und alle Gruppen $K$ zwischen $H$ und $G$ zu summieren hat.

Wegen $(K)=(H)(K/H)$ läßt sich hierfür

$$\varrho(z)=\sum_{H}{}'(H)^{-z} f(z;G/H) \tag{40}$$

schreiben mit

$$f(z,G)=\sum_{L\subseteq G} \mu(L)\,(L)^{-z}, \tag{41}$$

wobei man über alle Untergruppen $L$ von $G$ zu summieren hat. Wir haben nur noch zu zeigen, daß $f(z,G)$ der bei (20), (21) definierten Funktion $\psi(z,G)$ gleich ist, denn dann stimmt (40) mit (22) überein.

Haben $P_1,\ldots,P_s$ die Bedeutung wie in (21), so folgt aus (41) und der Definition von $\mu(G)$ leicht

$$f(z,G)=\prod_{i=1}^{s} f(z,P_i).$$

Dies mit (21) verglichen sehen wir, daß es genügt $f(z,P)=\psi(z,P)$ zu zeigen, wobei $P$ eine endliche Abelsche $p$-Gruppe ist. Wir bezeichnen mit $Q$ die größte, in $P$ enthaltene elementare Gruppe. Aus (41) und der Definition von $\mu$ folgt

$$f(z,P)=\sum_{L\subseteq Q} \mu(L)\,(L)^{-z}=f(z,Q). \tag{42}$$

Wir setzen

$$(Q)=p^t \qquad (t\geqq 1).$$

Dann ist $t$ zugleich die Anzahl der Invarianten von $P$. Wir setzen zur Abkürzung

$$\begin{bmatrix} k\\ l\end{bmatrix}=\frac{(p^k-1)(p^{k-1}-1)\cdots(p^{k-l+1}-1)}{(p-1)(p^2-1)\cdots(p^l-1)} \qquad (0\leqq l\leqq k)$$

und erweitern diese Definition durch die Vereinbarung $\begin{bmatrix} k\\ l\end{bmatrix}=0$ für $l<0$ und für $l>k$. Bekanntlich ist $\begin{bmatrix} t\\ l\end{bmatrix}$ die Anzahl der Untergruppen $p^l$-ter Ordnung von $Q$ $(l=0,\ldots,t)$. Für jede solche Gruppe $L$ gilt $\mu(L)=(-1)^l\,p^{\binom{l}{2}}$, somit

folgt aus (42)

$$f(z, P) = \sum_{l=0}^{t} (-1)^l p^{\binom{l}{2}} \begin{bmatrix} t \\ l \end{bmatrix} p^{-lz}.$$

Wir bezeichnen die rechte Seite kurz mit $g(z, t)$. Es genügt zu zeigen, daß $g(z, t)$ für $t = 0, 1, \ldots$ mit der rechten Seite von (20) übereinstimmt. Für $t = 0$ ist das klar, da $g(z, 0) = 1$ ist. Da ferner der Summand für $l < 0$ und für $l > t$ verschwindet, so gilt

$$g(z, t) = \sum_{l} (-1)^l p^{\binom{l}{2}} \begin{bmatrix} t \\ l \end{bmatrix} p^{-lz},$$

wobei man über alle $l = 0, \pm 1, \pm 2, \ldots$ summieren darf. Wegen

$$\begin{bmatrix} t+1 \\ l \end{bmatrix} = p^l \begin{bmatrix} t \\ l \end{bmatrix} + \begin{bmatrix} t \\ l-1 \end{bmatrix}$$

bekommt man

$$g(z+1, t+1) = \sum_{l} (-1)^l p^{\binom{l}{2}} \begin{bmatrix} t \\ l \end{bmatrix} p^{-lz} + \sum_{l} (-1)^l p^{\binom{l}{2}} \begin{bmatrix} t \\ l-1 \end{bmatrix} p^{-lz-l} \quad (t \geqq 0).$$

Wird im zweiten Summand $l+1$ statt $l$ geschrieben, so erkennt man, daß

$$g(z+1, t+1) = (1 - p^{-z-1}) g(z, t) \qquad (t \geqq 0)$$

ist. Dies bestätigt (20) mit vollständiger Induktion nach $t$, womit Satz 8 bewiesen ist.

Bezüglich des Satzes 9 beweisen wir zuerst die linke Seite von (28). Diese ist eine direkte Folgerung aus Satz 8, da $\psi(z, G)$ nach (20), (21) für $z = 1, 2, \ldots$ (sogar für $z = 0, 1, \ldots$) nichtnegativ ist. Aus der linken Seite von (28) folgt wegen Satz 5 auch die rechte Seite. Somit haben wir Satz 9 bewiesen.

Zum Beweis des Zusatzes nehmen wir Satz 11 vorweg. Aus diesem folgt

$$\varrho(z; A_1, \ldots, A_n) \leqq \varrho(z; A_i) \qquad (z = 1, 2, \ldots; \; i = 1, \ldots, n).$$

Die rechte Seite ist nach (1) gleich $1 - (A_i)^{-z}$. Somit haben wir den Zusatz von Satz 9 bewiesen.

Die Richtigkeit von Satz 10 (und die des Korollars) folgt aus Satz 8, wie wir es in der Bemerkung 6 schon auseinandergesetzt haben.

Satz 11 folgt sofort aus Satz 4 und der ersten Hälfte von Satz 9.

Um Satz 12 zu beweisen bemerken wir zunächst, daß nach (20)

$$\psi(0, P) = 0, \quad \psi'(0, P) = \log p \prod_{l=1}^{t-1} (1 - p^l) \qquad (t \geqq 1) \tag{43}$$

gilt. Aus $(43_1)$ folgt nach (21), daß $\psi'(0, G)$ verschwindet, wenn $(G)$ durch mindestens zwei verschiedene Primzahlen teilbar (d. h. $s \geqq 2$) ist. Auch gilt $\psi(0, G) = 0$, wenn $(G) \neq 1$ ist. Differenziert man also (1) und (22) an der

Stelle $z=0$, so folgt aus $(43_2)$ nach Gleichsetzung der rechten Seiten

$$-\sum_{\mathfrak{M}\subseteq\mathfrak{N}}(-1)^{(\mathfrak{M})}\log(A_{\mathfrak{M}})=\sum_{H}{}'\log p\prod_{l=1}^{t-1}(1-p^l),$$

wobei man über die Untergruppen $H(\neq G)$ von $G$ zu summieren hat, die kein $A_i$ enthalten und für die die Ordnung $(G/H)$ eine Primzahlpotenz $p^t$ $(t\geqq 1)$ ist (so daß also $p$ und $t$ sich von Glied zu Glied ändern). Hieraus folgt leicht die Gleichheit der rechten Seiten von (32), (33), womit wir Satz 12 bewiesen haben.

## § 4. Schlußbemerkungen

Wir wollen noch einige kurze Bemerkungen über unsere Zetafunktionen machen. Zunächst betrachten wir wieder den endlich-Abelschen Fall. Wie wir schon in der Einleitung bemerkt haben, halten wir den auf dem Hauptsatz (Satz 8) beruhenden Trägheitssatz (Satz 9) für das wichtigste Ergebnis der von uns in den §§ 2, 3 entwickelten „Theorie der Zetafunktionen für endliche Abelsche Gruppen". Über diese Theorie haben wir schon bemerkt, daß sie vom Fundamentalsatz der endlichen Abelschen Gruppen unabhängig ist, was wir jetzt bezüglich der Sätze 8—12 klarmachen wollen (für die Sätze 1—7 ist das ohnehin schon klar). Wohl bedeutet $t$ in (20) die Anzahl der Invarianten der $p$-Gruppe $P$ und diese Invarianten selbst bilden einen vom Fundamentalsatz abhängigen Begriff, aber $t$ läßt sich auch unabhängig vom Fundamentalsatz durch $p^t=(Q)$ definieren, wobei $Q$ die Gruppe derjenigen Elemente von $P$ bedeutet, die von der Ordnung $\leqq p$ sind. Da ferner die (eindeutige) Zerlegung von $G$ in seine Primärkomponenten $P_1,\ldots,P_s$ nichts mit dem Fundamentalsatz zu tun hat, so läßt sich auch die Funktion $\psi(z,G)$ in (21) vom Fundamentalsatz unabhängig definieren. Da endlich der Beweis des Satzes von DELSARTE [1] über die verallgemeinerte Umkehrungsformel von Möbius keinen unentbehrlichen Gebrauch vom Fundamentalsatz macht, so gilt ähnliches über den Beweis von unserem Satz 8. Es sind also auch die Sätze 9 bis 12, als Folgerungen aus Satz 8, vom Fundamentalsatz ebenfalls unabhängig. Diese Bemerkung ist auch deshalb interessant, weil eine ähnliche Unabhängigkeit auch der Satz von HAJÓS aufzeigt (vgl. [3]).

Es ist bemerkenswert, daß man zum Beweis vom Satz 9 nicht des vollen Inhalts vom Satz 8 bedarf. Und zwar genügt es hierzu Satz 8 nur für elementare $p$-Gruppen zu beweisen, woraus für denselben Spezialfall auch die linke Seite von (28) folgt. Hieraus entsteht dann der allgemeine Fall auf Grund der elementaren Sätze 3, 6 durch eine leichte Induktion.

Die Sätze 1 bis 7 lassen sich mit einigen Änderungen auf den Fall von beliebigen endlichen Gruppen und Ringen verallgemeinern — was wir dem Leser überlassen können — mit den Sätzen 8 bis 12 geht das aber

nicht mehr. (Vgl. den Anfang der Bemerkung 5 im § 2.) Überhaupt scheinen uns die Zetafunktionen für ganz beliebige Strukturen von sehr kompliziertem Verhalten zu sein.

Dagegen bieten die Zetafunktionen für die Abelschen Torsionsgruppen gewiß einen sehr interessanten Fall. (Eine Gruppe nennen wir eine Torsionsgruppe, wenn ihre Elemente von endlicher Ordnung sind.) Bezeichnet $G$ eine solche (unendliche) Gruppe, so ist es eine Grundaufgabe zu untersuchen, für welche $z(Rz \geqq 1)$ ein

$$\varrho(z) = \varrho(z; A_1, A_2, \ldots) \tag{44}$$

existiert, wobei $A_1, A_2, \ldots$ ein abzählbar unendliches System von (endlichen) Untergruppen von $G$ ist.

Diesbezüglich zeigen wir hier nur, daß (44) jedenfalls für alle $z = 1, 2, \ldots$ existiert. Für jedes solche $z$ ist nämlich die Folge

$$\varrho(z; A_1, \ldots, A_n) \qquad (n = 1, 2, \ldots)$$

nach den Sätzen 9, 11 beschränkt und monoton, weshalb sie konvergent ist. Wir haben noch zu zeigen, daß ihr Grenzwert von der Reihenfolge der $A_1, A_2, \ldots$ unabhängig ist. Zu diesem Zweck bezeichne $B_1, B_2, \ldots$ eine andere Anordnung von $A_1, A_2, \ldots$. Zu jedem $n$ gibt es dann ein $m$, so daß alle $A_1, \ldots, A_n$ unter den $B_1, \ldots, B_m$ vorkommen. Nach Satz 11 gilt dann

$$\varrho(z; A_1, \ldots, A_n) \geqq \varrho(z; B_1, \ldots, B_m),$$

woraus

$$\lim_{n \to \infty} \varrho(z; A_1, \ldots, A_n) \geqq \lim_{n \to \infty} \varrho(z; B_1, \ldots, B_n)$$

folgt. Aus Symmetriegründen gilt dies auch mit „$\leqq$", also auch mit „$=$". Das beendet den Beweis, womit wir die Existenz von (44) für $z = 1, 2, \ldots$ ausgewiesen haben. Auch folgt aus diesem Beweis, daß Satz 9 nebst Zusatz für (44) seine Gültigkeit behält, d. h.

$$0 \leqq \varrho(z; A_1, A_2, \ldots) \leqq 1 - (\min((A_1), (A_2), \ldots)^{-z} (<1) \qquad (z = 1, 2, \ldots)$$

ist.

Im Fall unendlicher Abelscher Gruppen haben wir (44) für die (reellen oder komplexen) Stellen $z \neq 1, 2, \ldots$ bisher gar nicht allgemein untersucht, weshalb wir über das etwaige analytische Verhalten der entsprechenden Zetafunktionen $\zeta(z) = \varrho(z; A_1, A_2, \ldots)^{-1}$ noch nicht orientiert sind. Die Untersuchung scheint für die (unendlichen) Abelschen Torsionsgruppen mit lauter endlichen $p$-Untergruppen aussichtsvoll zu sein. Allerdings gibt es unendliche Abelsche Gruppen, für die die Zetafunktionen sich trivial verhalten. So ist es z. B. mit der Prüferschen Gruppe $C(p^\infty)$ ($p$-Primzahl) definiert durch die Erzeugenden $\alpha_1, \alpha_2, \ldots$ und die zwischen ihnen bestehenden Gleichungen

$$\alpha_1^p = 1, \ \alpha_{i+1}^p = \alpha_i, \ \alpha_i \alpha_k = \alpha_k \alpha_i \qquad (i, k = 1, 2, \ldots).$$

Da jetzt die sämtlichen echten Untergruppen die Kette

$$\{\alpha_1\} \subset \{\alpha_2\} \subset \cdots$$

bilden, so folgt aus Satz 2 leicht, daß in diesem Fall stets

$$\varrho(z; A_1, A_2, \ldots) = \varrho(z; A_i) = 1 - (A_i)^{-z}$$

gilt, wobei $A_i$ das minimale unter den $A_1, A_2, \ldots$ bezeichnet. Dieses äußerst einfache Benehmen der Zetafunktionen für die Gruppen $C(p^\infty)$ steht im Einklang damit, daß diese Gruppen sonst auch in vieler Hinsicht eine Sonderstellung in der Gruppentheorie einnehmen. Soviel steht fest, daß die Zetafunktionen sehr empfindlich gegen die Kompliziertheit der zugrunde gelegten algebraischen Struktur reagiert.

*(Eingegangen am 22. Dezember 1954.)*

## Literaturverzeichnis

[1] S. Delsarte, Fonctions de Möbius sur les groupes Abelian finis, *Annals of Math.*, **49** (1948), S. 600—609.

[2] G. Hajós, Über einfache und mehrfache Bedeckung des $n$-dimensionalen Raumes mit einem Würfelgitter, *Math. Zeitschr.*, **47** (1942), S. 427—467.

[3] L. Rédei, Die gruppentheoretischen Zetafunktionen und der Satz von Hajós, *Acta Math. Acad. Sci. Hung.*, **6** (1955) (im Erscheinen).

## $\zeta$-ФУНКЦИИ В АЛГЕБРЕ

Л. Редеи (Сегед)

(Резюме)

Пусть $M$ — произвольное множество; обозначим через $(M)$ $(=0, 1, 2, \ldots, \infty)$ число элементов от $M$. Итак, если $M$ является (алгебраической) структурой (группой, кольцом, полугруппой и т. п.), то $(M)$ означает порядок от $M$.

Пусть $n$ — неотрицательное целое число. Обозначим через $\mathfrak{N}$ множество чисел $1, \ldots, n$. (Если $n=0$, то $\mathfrak{N}=0$ является пустым множеством.)

Пусть дана теперь структура $S$, и посмотрим систему $A_1, \ldots, A_n$ ее подструктур. (Если $S$ группа или кольцо и. т. п., то каждое из множеств $A_i$ является также группой, или кольцом и. т. п.) При любом $\mathfrak{M} (\subseteq \mathfrak{N})$ обозначим через $A_{\mathfrak{M}}$ подструктуру из $S$ порожденную элементами подструктур $A_i$ $(i \in \mathfrak{M})$. Рассмотрим систему

$$(1) \qquad (A_{\mathfrak{M}}) \qquad (0 \subset \mathfrak{M} \subseteq \mathfrak{N}).$$

(Она не легко обозрима даже в том частном случае, когда $S$ конечная абелева группа и $n \geqq 3$.)

Чтобы получить сведения о системе (1), определим с помощью комплексной переменной $z$ для $Rz > 0$ функцию

$$(2) \qquad \varrho(z) = \varrho(z; A_1, \ldots, A_n) = \sum_{\mathfrak{M} \subseteq \mathfrak{N}} (-1)^{(\mathfrak{M})} (A_{\mathfrak{M}}) \quad .$$

где $Rz$ означает вещественную часть $z$ и в случае $\mathfrak{M}=0$ следует считать $(A_{\mathfrak{M}})=1$. (Конечно, можно пренебрегать членами соответствующими случаями $(A_{\mathfrak{M}})=\infty$, и для положительных $c$ следует считать $c^{-z}=e^{-z\log c}$, где $\log c$ означает вещественное значение логарифма.) Будем называть функции

$$\zeta(z)=\zeta(z;A_1,\ldots,A_n)=\varrho(z)^{-1}=\varrho(z;A_1,\ldots,A_n)^{-1} \tag{3}$$

$\zeta$-функциями (в алгебре). Распространим наши определения и на случай бесконечных (счетных) систем подструктур $A_1, A_2, \ldots$ так, чтобы было

$$\varrho(z;A_1,A_2,\ldots)=\lim_{n\to\infty}\varrho(z;A_1,\ldots,A_n), \tag{4}$$

если только правая сторона существует и ее значение независимо от порядка членов в последовательности $A_1, A_2, \ldots$. С помощью аналитического продолжения можно распространять определение в некоторых случаях на всю плоскость. Можно охватывать и структуры с операторами, ограничиваясь подструктурами, допустимыми относительно операторов. Если $S$ группа или кольцо, то можно определение „дуализировать", взяв вместо $A_{\mathfrak{M}}$ пересечение подструктур $A_i$ $(i\subseteq\mathfrak{M})$, и вместо порядкового числа $(A_{\mathfrak{M}})$ индекс $(S:A_{\mathfrak{M}})$ („двойственная $\zeta$-функция"). (Если однако $S$ конечная абелева группа, то двойственные $\zeta$-функции в совокупности совпадают с „обыкновенными" $\zeta$-функциями.)

Если, в частности, $A_i$ $(i=1,2,\ldots)$ есть группа порядка $p_i$, где $p_i$ означает $i$-ое простое число, а $S$ является прямым произведением групп $A_1, A_2, \ldots$, то очевидно

$$\zeta(z;A_1,A_2,\ldots)=\varrho(z;A_1,A_2,\ldots)^{-1}\prod_{i=1}^{\infty}(1-p_i^{-z})^{-1},$$

а это есть как раз $\zeta$-функция Римана. Если $S$ бесконечная циклическая группа, а $A_i$ $(i=1,2,\ldots)$ ее подгруппы индексов $p_i$, то соответствующая двойственная $\zeta$-функция есть опять $\zeta$-функция Римана. Если же $\mathsf{P}$ кольцо целых чисел алгебраического числового тела конечной степени, а $A_1, A_2, \ldots$ все простые идеалы $\mathsf{P}$ (иными словами: все максимальные допустимые подмодули операторного модуля $\mathsf{P}^+$, где операция означает умножение на элементы $\mathsf{P}$), то соответствующая двойственная $\zeta$-фуикция есть $\zeta$-функция Дедекинда, принадлежащей $\mathsf{P}$.

Автором был до сих пор более подробно рассмотрен только случай, когда $S=G$ есть конечная абелева группа. Кроме нескольких почти тривиальных соотношений (теоремы 1—7) имеет место „основная теорема $\zeta$-функций конечных абелевых групп":

Имеет место соотношение

$$\varrho(z;A_1,\ldots,A_n)=\sum_H (H)^{-z}\psi(G/H), \tag{5}$$

где $H$ пробегает подгруппы $G$ не содержающие ни одну из подгрупп $A_1,\ldots,A_n$, $G/H$ означает факторгруппу а функция $\psi$ определена следующим образом: Если $P$ конечная абелева $p$-группа с конечным числом $t$ инвариантов, то

$$\psi(P)=(1-p^{-z})(1-p^{1-z})\cdots(1-p^{t-s-z}),$$

если же $P_1,\ldots,P_k$ все различные силовские подгруппы конечной абелевой группы, то

$$\psi(P_1\cdots P_k)=\psi(P_1)\cdots\psi(P_k).$$

Наиболее важным следствием является „теорема инерции конечных абелевых групп", соответственно которой всегда имеет место

$$0\leqq\varrho(z;A_1,\ldots,A_n)<1 \qquad (z=1,2,\ldots). \tag{6}$$

Эта теорема является оптимальной с трех точек зрения. Во первых, если $z\,(\geqq 1)$ вещественное нецелое число, то ни одно из неравенств

$$c<\varrho(z;A_1,\ldots,A_n)<C$$

не выполняется всегда, как бы ни задать вещественные числа $c, C$. Во вторых для любого целого положительного числа $z$ на левой стороне (6) имеет место равенство во бесконечном числе (нетривиальных) случаев. Во третьях, если $A_1, \ldots, A_n$ подгруппы конечной некоммутативной группы, то левая сторона (6) вообще несправедлива, даже

$$c < \varrho(z; A_1, \ldots, A_n) \qquad (z = 1, 2, \ldots)$$

не выполняется ни при каком $c$.

Дальнейшим следствием из (5) является (в случае конечных абелевых групп) „свойство монотонности"

$$\varrho(z; A_1, \ldots, A_n) \leqq \varrho(z; A_1, \ldots, A_{n-1}) \qquad (z = 1, 2, \ldots).$$

Можно ожидать интересные результаты и от исследования (бесконечных) периодических абелевых групп. Для них $\varrho(z; A_1, A_2, \ldots)$ существует при $z = 1, 2, \ldots$ и в аналогии с (6) имеет место

$$0 \leqq \varrho(z; A_1, A_2, \ldots) < 1 \qquad (z = 1, 2, \ldots).$$

# NEUER BEWEIS DES HAJÓSSCHEN SATZES ÜBER DIE ENDLICHEN ABELSCHEN GRUPPEN

Von

LADISLAUS RÉDEI (Szeged), Mitglied der Akademie

*Herrn Prof. Dr.* LADISLAUS KALMÁR *am 50. Geburtstage mit Freundschaft zugeeignet*

*Durchgängige Bezeichnungen:*

$G$ eine endliche Abelsche Gruppe mit mindestens zwei Elementen.

$\varepsilon$ Einselement von $G$.

$\alpha, \beta, \ldots$ (mit oder ohne Indizes) Elemente von $G$.

$a, b, \ldots$ (mit oder ohne Indizes) ganze Zahlen.

$p, q, r$ (mit oder ohne Indizes) positive Primzahlen.

$\{\ldots\}$ die durch die eingeklammerten Elemente erzeugte Untergruppe von $G$.

$o(\ldots)$ die Ordnung eines Gruppenelementes.

$O(\ldots)$ die Ordnung einer Gruppe.

$O\{\ldots\} = O(\{\ldots\})$.

$A$ (mit oder ohne Indizes) ein Komplex (d. h. eine Untermenge) von $G$. Unter dem Produkt $A_1 \ldots A_k$ der Komplexe $A_1, \ldots, A_k$ verstehen wir wie üblich den Komplex aller verschiedenen $\alpha_1 \ldots \alpha_k$ ($\alpha_i \in A_i$; $i = 1, \ldots, k$). Erscheint in dieser Form jedes Element von $A_1 \ldots A_k$ nur einmal, so nennen wir dieses Komplexprodukt *direkt*. Bei uns werden ausschließlich direkte Komplexprodukte vorkommen, auch wenn wir es nicht ausdrücklich betonen.

$(\alpha)_e$ der Komplex mit den Elementen $\varepsilon, \alpha, \ldots, \alpha^{e-1}$ ($o(\alpha) \geqq e \geqq 2$). Wir nennen $(\alpha)_e$ ein *Simplex* (von $G$). Dies ist dann und nur dann eine (zyklische) Gruppe, wenn $o(\alpha) = e$ ist. Insbesondere nennen wir $(\alpha)_p$ ein *Primsimplex*. Gilt $G = (\alpha_1)_{e_1} \cdots (\alpha_n)_{e_n}$ und ist die rechte Seite ein direktes Produkt, so nennen wir sie eine *Hajóssche Zerlegung* (von $G$). Diese nennen wir im Fall $e_i = p_i$ ($i = 1, \ldots, n$) *prim*.

## § 1. Einleitung

Der berühmte, vor 14 Jahren entdeckte Satz von HAJÓS [2][1] (s. auch HAJÓS [1]) lautet so:

SATZ 1. *In jeder Hajósschen Zerlegung von G ist mindestens ein Faktor eine Gruppe.*

Diesem Satz verleihen neben seiner Wichtigkeit auch die unerwarteten Schwierigkeiten des Beweises einen besonderen Reiz. Wohl ist nämlich der Satz dem Wortlaut nach eine der einfachsten in der Gruppentheorie, auch bieten sich zum Beweis viele trivial erscheinende Wege an, zur peinlichen Überraschung zeigen sich aber die meisten von ihnen sehr bald als Irrwege, die nur da sind, um einem zu erschweren, einen richtigen Weg zum Beweis zu finden.

HAJÓS überwand die Schwierigkeiten mit einer genialen Idee, die nach Heranziehen des Ringes $\mathfrak{J}(G)$ von $G$ über dem Ring $\mathfrak{J}$ der ganzen Zahlen hauptsächlich in einer feinen Nullteileruntersuchung bestand. Sein glänzender Beweis ist bis heute der einzige geblieben, doch wurde dieser durch RÉDEI [5] und noch weiter durch SZELE [9] bedeutend vereinfacht. (S. auch das Buch von RÉDEI [7], wo dieser vereinfachte Beweis möglichst kurz ausgearbeitet ist.) Wir bemerken, daß wegen der Abstraktheit des von tiefen Gedanken erfüllten Beweises von HAJÓS sein Satz auch noch heute sehr geheimnisvoll ist.

Ich habe früher (vgl. RÉDEI [6]) den Satz rein gruppentheoretisch, prinzipiell sehr einfach aber in der Ausführung ziemlich kompliziert auf den Spezialfall der $p$-Gruppen reduziert, konnte jedoch damals auf diesem Wege nicht zu einem vollständigen Beweis kommen. Es schien deshalb, daß die eigentlichen Schwierigkeiten des Satzes in diesem Spezialfall konzentriert sind, welche Meinung auch im Beweis von Hajós ihre Unterstützung findet, da dieser sich für $p$-Gruppen nicht kürzer gestaltet.

Nach vielem Mißerfolg, aber von den Schönheiten des Satzes immer neu angeregt fand ich endlich für den Fall von $p$-Gruppen einen überraschend einfachen Beweis. Vollständigkeitshalber führe ich den Beweis auch für nicht-$p$-Gruppen aus; dieser Teil des Beweises ist im wesentlichen eine Reproduktion meiner vorigen Arbeit [6], doch mit einigen Vereinfachungen.

Am Ende meines Beweises für $p$-Gruppen benütze ich einen kurzen ringtheoretischen Schluß im auch von HAJÓS verwendeten Gruppenring $\mathfrak{J}(G)$, sonst ist mein Beweis gruppentheoretisch. Als ich das Herrn SZELE erzählte, hat er bemerkt, daß sich dieser ringtheoretische Schluß wahrscheinlich durch einen gruppentheoretischen ersetzen läßt. Das ist wenn auch etwas kompliziert in der Tat möglich, das ich im § 5 zeigen werde. Es ist die prinzipielle Wichtigkeit zu betonen, daß somit für den Satz von HAJÓS durch die

[1] [ ] bezieht sich auf das Literaturverzeichnis am Ende unserer Arbeit.

wertvolle Mitwirkung von Herrn SZELE auch ein rein gruppentheoretischer Beweis entstanden ist.

Der Satz von HAJÓS brachte zurzeit auch die Überraschung mit, daß ihn HAJÓS unabhängig vom Fundamentalsatze der endlichen Abelschen Gruppen bewiesen hat und dieser nicht geeignet ist seinen Beweis zu verkürzen. Dasselbe bezieht sich auch auf meinen Beweis, weshalb beide Sätze mit großer Wahrscheinlichkeit voneinander völlig unabhängig sind. Das ist sehr merkwürdig, da es natürlich möglich ist, mit Hilfe des Fundamentalsatzes den Satz von HAJÓS in Angriff zu nehmen; desto enttäuschender ist seine Machtlosigkeit zum Beweis.

Bezüglich des Verhältnisses beider Sätze zueinander soll auch bemerkt werden, daß sich der Satz von HAJÓS als eine Umkehrung des Fundamentalsatzes ansehen läßt, da man diesen folgenderweise aussprechen kann: Jede endliche Abelsche Gruppe hat eine Hajóssche Zerlegung, in der alle Faktoren Gruppen sind.

Nach diesen Bemerkungen liegt vor uns eine Synthese von zwei gleich wichtigen, einander ergänzenden und (mit großer Wahrscheinlichkeit) voneinander unabhängigen Sätzen, nämlich des Fundamentalsatzes und des Satzes von HAJÓS.

Es ist auch eine merkwürdige Gegenüberstellung beider Sätze, daß der Fundamentalsatz sich von $p$-Gruppen auf nicht-$p$-Gruppen trivial verallgemeinern läßt, dagegen in meinem Beweis eben der Fall von $p$-Gruppen sehr einfach und die Verallgemeinerung auf nicht-$p$-Gruppen kompliziert ist. Vielleicht läßt auch diese Unverhältnismäßigkeit hoffen, daß der zweite Teil meines Beweises vereinfachbar ist.

Ein wesentliches Moment ist es, daß wir den Hajósschen Satz in seiner von uns bald nach seinem Entstehen entdeckten „scharfen" Form beweisen werden. (Vgl. RÉDEI [3]. Es ist ein eigenartiger Zug vom ganzen Problemkreis des Minkowski—Hajósschen Satzes, daß auch die beiden, dem Satz von HAJÓS äquivalenten Minkowskischen Vermutungen je eine „scharfe" Formulierung zulassen. Vgl. RÉDEI [4], wo diese drei Sätze in „schwacher" und „scharfer" Form angeführt sind.)

Um die scharfe Form des Satzes von HAJÓS aussprechen zu können, betrachten wir *Systeme*

$$\begin{pmatrix} \alpha_1 & \dots & \alpha_n \\ e_1 & \dots & e_n \end{pmatrix} \qquad (e_1, \dots, e_n \geqq 2;\ n \geqq 1). \tag{1}$$

Wir nennen (1) ein *geordnetes Normalsystem* (für $\{\alpha_1, \dots, \alpha_n\}$), wenn

$$O\{\alpha_1, \dots, \alpha_k\} = e_1 \dots e_k \qquad (k = 1, \dots, n)$$

ist. Mit anderen Worten bedeutet das, daß zur Untergruppenkette

$$\varepsilon \subset \{\alpha_1\} \subset \{\alpha_1, \alpha_2\} \subset \dots \subset \{\alpha_1, \dots, \alpha_n\}$$

die Indizesfolge $e_1, \dots, e_n$ gehört. (Für Primzahlen $e_1, \dots, e_n$ handelt es sich

um Kompositionsreihen.) Unter einem *Normalsystem* schlechthin verstehen wir ein System (1), das durch eine Vertauschung der „Spalten“ $\alpha_i, e_i$ $(i=1,\ldots,n)$ in ein geordnetes Normalsystem überführbar ist. (Statt „Normalsystem“ sagen wir auch „normales System“.) Nunmehr verstehen wir unter dem *scharfen Hajósschen Satz* den folgenden:

Satz 2. *Dann und nur dann ist*

$$G=(\alpha_1)_{e_1}\cdots(\alpha_n)_{e_n} \tag{2}$$

*eine Hajóssche Zerlegung, wenn* (1) *ein Normalsystem für G ist.*

Dieser Satz ist in der Tat eine Verschärfung vom Satz 1, denn in einem Normalsystem (1) muß nach seiner Definition für mindestens ein $i$ die Gleichung $o(\alpha_i)=O\{\alpha_i\}=e_i$ erfüllt, d. h. der Faktor $(\alpha_i)_{e_i}$ eine Gruppe sein.

Gleich hier beweisen wir den Teil „dann“ vom Satz 2. Im Fall eines Normalsystems für $G$ sind die Produkte

$$\alpha_1^{k_1}\cdots\alpha_n^{k_n} \quad (k_i=0,\ldots,e_i-1\,;\ i=1,\ldots,n)$$

offenbar die sämtlichen verschiedenen Elemente von $G$, was eben besagt, daß (2) eine Hajóssche Zerlegung von $G$ ist.

Da hiernach der Teil „dann“ vom Satz 2 eine Trivialität ist, somit das Gewicht von ihm im Teil „nur dann“ liegt, den wir allein noch zu beweisen haben, so werden wir im folgenden „Satz 2“ oft mit seinem Teil „nur dann“ identifizieren.

Übrigens werden wir sehen, daß Satz 2 eine fast triviale Folgerung vom Satz 1 ist (vgl. Hilfssatz 1 nebst Korollar im § 2).

Natürlich ist Satz 2 insofern Satz 1 überlegen, als er die Hajósschen Zerlegungen vollständig charakterisiert, während Satz 1 nur eine notwendige Bedingung für sie ausspricht. Diese Überlegenheit macht Satz 2 für einen induktiven Beweis besonders geeignet.

Wir bemerken noch, daß die leichte aber wichtige Reduktion von Hajós auf den Fall von Primsimplexen auch in unserem Beweis unentbehrlich wird. Sonst haben beide Beweise keine Berührungspunkte.[2]

## § 2. Reduktion des Beweises

Hilfssatz 1. *Gilt Satz 1 für eine gegebene Hajóssche Zerlegung* (2) *von G und gilt Satz 2 für die Gruppen von kleinerer Ordnung als G, so gilt Satz 2 für die gegebene Zerlegung* (2).

Korollar. *Hieraus folgt die Äquivalenz der Sätze 1, 2.*

[2] Bezüglich anderer neuerer Forschungen im Problemkreis des Satzes von Hajós s. man Rédei [8].

Zum Beweis legen wir uns eine Hajóssche Zerlegung (2) von $G$ vor, in der z. B. $(\alpha_1)_{e_1}$ eine Gruppe ist. Dann ist diese gleich $\{\alpha_1\}$, ferner gilt $O\{\alpha_1\}=o(\alpha_1)=e_1$. Aus (2) folgt $G=\{\alpha_1, \ldots, \alpha_n\}$. Nun ist im Fall $n=1$ die Behauptung klar. Im Fall $n \geqq 2$ gewinnen wir aus dem Gesagten, daß

$$(\alpha_2\{\alpha_1\})_{e_2}\cdots(\alpha_n\{\alpha_1\})_{e_n}$$

eine Hajóssche Zerlegung der Faktorgruppe $G/\{\alpha_1\}$ ist. Wegen der Voraussetzung ist also

$$\begin{pmatrix} \alpha_2\{\alpha_1\} & \cdots & \alpha_n\{\alpha_1\} \\ e_2 & \cdots & e_n \end{pmatrix}$$

ein Normalsystem (für $G/\{\alpha_1\}$). Hieraus und aus $O\{\alpha_1\}=e_1$ folgt, daß (1) ein Normalsystem für $G$ ist, was wir zu beweisen hatten.

Hilfssatz 2. (Hajós) *Ist Satz 1 für die primen Hajósschen Zerlegungen von G richtig, so ist er für alle Hajósschen Zerlegungen von G richtig.*

Denn legen wir uns eine Hajóssche Zerlegung (2) vor. Diese läßt sich mit wiederholter Anvendung der Formel

$$(\alpha)_{ef}=(\alpha)_e(\alpha^e)_f \qquad (o(\alpha) \geqq ef;\ e, f \geqq 2)$$

in eine prime Hajóssche Zerlegung verwandeln. Zum Beweis vom Hilfssatz 2 genügt es also folgendes zu zeigen. Gilt eine direkte Produktzerlegung

$$(\varrho)_e=AB \qquad (A \neq \varepsilon),$$

wo $A$ eine Gruppe ist, so ist auch $(\varrho)_e$ eine Gruppe. Um dies zu beweisen nehmen wir ein Element $\alpha(\neq\varepsilon)$ von $A$. Wegen $\alpha A=A$ besteht

$$\alpha(\varrho)_e=(\varrho)_e.$$

Hieraus folgt $\alpha=\varrho^k$ $(0<k<e)$, $\varrho^e=\alpha\varrho^{e-k}=\varrho^l$ $(0 \leqq l<e)$, $\varrho^{e-l}=\varepsilon$, also wegen $o(\varrho) \geqq e$ auch $o(\varrho)=e$, womit der Beweis beendet ist.

Aus den Hilfssätzen 1, 2 kommen wir zur folgenden:

Reduktion des Beweises. *Gilt Satz 1 für die primen Hajósschen Zerlegungen von G und gilt Satz 2 für die Gruppen von kleinerer Ordnung als G, so gilt letzterer auch für G.*

Aus Hilfssatz 2 folgt nämlich, daß Satz 1 für $G$ gilt. Weiter folgt hieraus nach Hilfssatz 1 in der Tat die Richtigkeit von Satz 2 für $G$.

## § 3. Beweis für $p$-Gruppen.

Lemma. *Ist für* $\alpha_1, \ldots, \alpha_n$ $(n \geqq 1)$

(3) $$O\{\alpha_1, \ldots, \alpha_n\}=p^n$$

*und gelten alle Ungleichungen*

(4) $$O\{\alpha_{i_1}, \ldots, \alpha_{i_k}\} \geqq p^k \qquad (1 \leqq i_1<\cdots<i_k \leqq n;\ 1 \leqq k \leqq n-1),$$

*so gibt es Potenzen $s_1, \ldots, s_n$ von $p$, so daß*

$$(5) \qquad \begin{pmatrix} \alpha_1^{s_1} & \cdots & \alpha_n^{s_n} \\ p & \cdots & p \end{pmatrix}$$

*ein Normalsystem (für $\{\alpha_1, \ldots, \alpha_n\}$) ist.*

Im Fall $n=1$ ist die Behauptung trivial. Im Fall $n \geqq 2$ machen wir die Induktionsannahme, daß die Behauptung für die Systeme $\alpha_1, \ldots, \alpha_n$ mit kleinerem $o(\alpha_1)\cdots o(\alpha_n)$ richtig ist. *Erstens* betrachten wir den Fall, daß in (4) mindestens einmal „$=$“ gilt. Dann dürfen wir mit passender Reihenfolge der $\alpha_1, \ldots, \alpha_n$ annehmen, daß für ein $m$ $(1 \leqq m \leqq n-1)$

$$(6) \qquad O\{\alpha_1, \ldots, \alpha_m\} = p^m$$

ist. Da ferner (4) für $m$ statt $n$ noch mehr gilt, so gibt es nach der Induktionsannahme Potenzen $s_1, \ldots, s_m$ von $p$, so daß

$$(7) \qquad \begin{pmatrix} \alpha_1^{s_1} & \cdots & \alpha_m^{s_m} \\ p & \cdots & p \end{pmatrix}$$

ein Normalsystem ist. Hieraus und aus (6) folgt auch

$$(8) \qquad \{\alpha_1, \ldots, \alpha_m\} = \{\alpha_1^{s_1}, \ldots, \alpha_m^{s_m}\}.$$

Wird andererseits

$$(9) \qquad \hat{\alpha} = \alpha\{\alpha_1, \ldots, \alpha_m\}$$

gesetzt, aufgefaßt als Element der Faktorgruppe $G/\{\alpha_1, \ldots, \alpha_m\}$, so folgen aus (3), (4), (6) offenbar

$$O\{\hat{\alpha}_{m+1}, \ldots, \hat{\alpha}_n\} = p^{n-m},$$
$$O\{\hat{\alpha}_{i_1}, \ldots, \hat{\alpha}_{i_k}\} = p^k \qquad (m+1 \leqq i_1 < \cdots < i_k \leqq n;\ 1 \leqq k \leqq n-m-1).$$

Da ferner stets $o(\hat{\alpha}) \mid o(\alpha)$ ist, so folgt hieraus wieder nach der Induktionsannahme die Existenz von Potenzen $s_{m+1}, \ldots, s_n$ von $p$, so daß

$$(10) \qquad \begin{pmatrix} \hat{\alpha}_{m+1}^{s_{m+1}} & \cdots & \hat{\alpha}_n^{s_n} \\ p & \cdots & p \end{pmatrix}$$

ein Normalsystem ist. Wegen (8), (9) setzen sich die Normalsysteme (7), (10) zum Normalsystem (5) zusammen, womit die Behauptung für diesen Fall bewiesen ist. *Zweitens* haben wir nur noch den Fall zu betrachten, daß (statt (4)) sogar

$$(11) \qquad O\{\alpha_{i_1}, \ldots, \alpha_{i_k}\} \geqq p^{k+1} \qquad (1 \leqq i_1 < \cdots < i_k \leqq n;\ 1 \leqq k \leqq n-1)$$

gilt. Insbesondere ist dann $O\{\alpha_2, \ldots, \alpha_n\} \geqq p^n$, also wegen (3)

$$O\{\alpha_1^p, \alpha_2, \ldots, \alpha_n\} = p^n.$$

Hiernach und nach (11) behalten (3), (4) ihre Gültigkeit bei Ersetzung von $\alpha_1$ durch $\alpha_1^p$. Wegen (4) gilt dabei $o(\alpha_1) \geqq p$, also $o(\alpha_1^p) < o(\alpha_1)$, somit folgt aus der Induktionsannahme mit Anwendung auf das System $\alpha_1^p, \alpha_2, \ldots, \alpha_n$ die Richtigkeit der Behauptung auch für diesen Fall (jetzt sogar mit $s_1 \geqq p$). Das beweist unser Lemma.

Nunmehr können wir Satz 2 für $p$-Gruppen beweisen. Nach der Reduktion im § 2 genügt es zu zeigen, daß in jeder Hajósschen Zerlegung

$$G = (\alpha_1)_p \dots (\alpha_n)_p \tag{12}$$

mindestens ein Faktor eine Gruppe ist.

Wegen der Annahme gilt (3), ferner ist jedes Teilprodukt von (12) ebenfalls direkt, weshalb auch (4) erfüllt ist. Nach dem Lemma gibt es also Potenzen $s_1, \dots, s_n$ von $p$, so daß (5) ein Normalsystem ist. Gilt dabei $s_1 = \dots = s_n = 1$, so sind wir mit dem Beweis fertig. Im anderen Falle dürfen wir annehmen, daß mit einem $m$

$$p | s_1, \dots, s_m; \; s_{m+1} = \dots = s_n = 1 \qquad (1 \leqq m \leqq n) \tag{13}$$

gilt. Da (5) ein Normalsystem ist, so folgt unter Berücksichtigung von (12) aus dem Teil „dann" vom Satz 2, daß auch

$$G = (\alpha_1^{s_1})_p \dots (\alpha_m^{s_m})_p (\alpha_{m+1})_p \dots (\alpha_n)_p \tag{14}$$

eine Hajóssche Zerlegung einer Gruppe ist. Wir werden in einen Widerspruch geraten, womit der Beweis beendet wird.

Einen Augenblick nehmen wir den im § 1 erwähnten Gruppenring $\mathfrak{J}(G)$ zu Hilfe. Gleichzeitig ändern wir unsere Bezeichnungen so ab, daß wir mit $(\alpha)_e$ das Element $\varepsilon + \alpha + \dots + \alpha^{e-1}$ von $\mathfrak{J}(G)$ bezeichnen, nach dem die rechte Seite von (12) ebenfalls ein Element von $\mathfrak{J}(G)$, und zwar die Summe der Elemente der Gruppe $G$ bedeutet. Da diese $\alpha_1$ enthält, so ist das $(\varepsilon - \alpha_1)$-fache der rechten Seite von (12) gleich 0. Wegen (13) und der Formel

$$(\varepsilon - \alpha)(\alpha)_e (\alpha^e)_f = \varepsilon - \alpha^{ef}$$

gilt also noch mehr

$$(\varepsilon - \alpha_1^{s_1}) \dots (\varepsilon - \alpha_m^{s_m})(\alpha_{m+1})_p \dots (\alpha_n)_p = 0.$$

Nach Ausmultiplizieren tritt links das Glied $\varepsilon$ auf, welches also gegen ein anderes Glied aufgehen muß. Das steht aber in einem Widerspruch damit, daß das Simplexprodukt (14) direkt ist, womit wir Satz 2 für $p$-Gruppen bewiesen haben.

## § 4. Beweis für nicht $p$-Gruppen

Hilfssatz 3. *Es gilt*

$$(\alpha_1)_{p_1} \dots (\alpha_n)_{p_n} = (\alpha_1^{t_1})_{p_1} \dots (\alpha_n^{t_n})_{p_n} \qquad (p_i \nmid t_i;\, i = 1, \dots, n), \tag{15}$$

*vorausgesetzt, daß die linke Seite eine Hajóssche Zerlegung einer Gruppe ist, ferner gilt dann*

$$p_i | o(\alpha_i) \qquad (i = 1, \dots, n).$$

Es genügt nämlich folgendes zu zeigen: Es gilt

$$(\alpha)_p A = (\alpha^t)_p A \qquad (p \nmid t), \tag{16}$$

wenn die linke Seite die Zerlegung einer Gruppe in ein direktes Produkt

und $\varepsilon \in A$ ist, ferner gilt dann $p|o(\alpha)$. Um dies zu beweisen berücksichtigen wir, daß $\alpha \in (\alpha)_p A$, also $\alpha(\alpha)_p A = (\alpha)_p A$ ist. Wegen $\alpha(\alpha)_p = \alpha, \ldots, \alpha^p$ folgt hieraus $\alpha^p A = A$, also allgemeiner

$$\alpha^{p^i} A = A \qquad (i = 0, \pm 1, \ldots).$$

Hiernach muß $p|o(\alpha)$ sein, denn im anderen Fall würde $\alpha A = A$ folgen, was aber wegen der Annahme unmöglich ist. Da ferner $\alpha^k A$ nur von der Restklasse $k \pmod p$ abhängt, so ist auch (16) richtig. Das beendet den Beweis vom Hilfssatz 3.

Hilfssatz 4. *Im Falle eines geordneten Normalsystems*

$$\begin{pmatrix} \alpha_1 & \ldots & \alpha_m \\ p & \ldots & p \end{pmatrix} \qquad (p^2 | o(\alpha_2), \ldots, o(\alpha_m);\ m \geqq 1)$$

*(für $\{\alpha_1, \ldots, \alpha_n\}$) läßt sich jedes Element $\alpha (\neq \varepsilon)$ von $\{\alpha_1, \ldots, \alpha_m\}$ in der Form*

$$\alpha = \alpha_1^{k_1} \ldots \alpha_m^{k_m} \qquad (p \nmid k_1;\ p \nmid k_i \text{ oder } k_i = 0;\ i = 2, \ldots, m)$$

*schreiben.*

Wir machen die Induktionsannahme, daß die Behauptung für die kleineren Werte von $m$ richtig ist. Aus der Voraussetzung folgt das Bestehen einer Gleichung

$$\alpha = \alpha_s^{l_s} \ldots \alpha_m^{l_m} \qquad (1 \leqq l_s \leqq p-1;\ 0 \leqq l_i \leqq p-1;\ i = s+1, \ldots, m).$$

Im Fall $s = 1$ sind wir mit dem Beweis fertig. Im Fall $s \geqq 2$ berücksichtigen wir, daß $\alpha_s^p$ wegen der Voraussetzung ein von $\varepsilon$ verschiedenes Element von $\{\alpha_1, \ldots, \alpha_{s-1}\}$ ist. Da ferner auch

$$\begin{pmatrix} \alpha_1 & \ldots & \alpha_{s-1} \\ p & \ldots & p \end{pmatrix}$$

ein geordnetes Normalsystem ist, so folgt aus der Induktionsannahme das Bestehen einer Gleichung

$$\alpha_s^p = \alpha_1^{k_1} \ldots \alpha_{s-1}^{k_{s-1}} \qquad (p \nmid k_1;\ p \nmid k_i \text{ oder } k_i = 0;\ i = 2, \ldots, s-1).$$

Aus beiden Gleichungen folgt

$$\alpha = \alpha_1^{k_1} \ldots \alpha_{s-1}^{k_{s-1}} \alpha_s^{-p+l_s} \alpha_{s+1}^{l_{s+1}} \ldots \alpha_m^{l_m},$$

womit Hilfssatz 4 bewiesen ist.

Hilfssatz 5. *Ist*

$$G = (\alpha_1)_{p_1} \ldots (\alpha_n)_{p_n} \tag{17}$$

*eine Hajóssche Zerlegung und sind*

$$\begin{pmatrix} \alpha_i & \alpha_{i_1} & \ldots & \alpha_{i_q} \\ p_i & p_i & \ldots & p_i \end{pmatrix}, \begin{pmatrix} \alpha_j & \alpha_{j_1} & \ldots & \alpha_{j_k} \\ p_j & p_j & \ldots & p_j \end{pmatrix} \qquad (i \neq j;\ 1 \leqq i, i_1, \ldots, j, j_1, \ldots, \leqq n) \tag{18}$$

*zwei geordnete Normalsysteme, wobei die in* (18) *vorkommenden $\alpha_{i_x}, \alpha_{j_y}$ von Primzahlpotenzordnung sind, so sind $\alpha_i, \alpha_{i_1}, \ldots, \alpha_j, \alpha_{j_1}, \ldots$ verschieden.*

Jedenfalls sind nämlich $\alpha_i, \alpha_{i_1}, \ldots$, desgleichen auch $\alpha_j, \alpha_{j_1}, \ldots$ untereinander verschieden. Ferner ist wegen der Voraussetzung kein Glied dieser Folgen dem ersten Glied der anderen Folge gleich. Wenn also die Behauptung falsch ist, so dürfen wir (nötigenfalls nach Streichung einiger hinterer Spalten der Systeme (18)) annehmen, daß die gesagten zwei Folgen mindestens zweigliedrig und ihre letzten Glieder $\alpha_{i_g} = \alpha_{j_h}$ gleich sind, sie aber sonst keine weiteren gleichen Glieder enthalten. Wegen $\alpha_{i_g} = \alpha_{j_h}$ müssen $p_i, p_j$ gleich sein. Wir setzen

$$p = p_i = p_j, \quad \alpha = \alpha_{i_g}^p = \alpha_{j_h}^p \ (\neq \varepsilon).$$

Da die Systeme (18) normal sind, so gilt

$$\alpha \in \{\alpha_i, \alpha_{i_1}, \ldots, \alpha_{i_{g-1}}\}, \ \{\alpha_j, \alpha_{j_1}, \ldots, \alpha_{j_{h-1}}\}.$$

Es ist auch klar, daß die $o(\alpha_i), o(\alpha_{i_x}), o(\alpha_j), o(\alpha_{j_y})$ Potenzen von $p$ sind. Hiernach können wir Hilfssatz 4 auf das Element $\alpha$ und beide geordneten Normalsysteme anwenden, die aus (18) nach Streichung der letzten Spalte entstehen. So kommen wir zu einer Gleichung von der Form

$$(\alpha =)\, \alpha_i^a \ldots \alpha_{i_{g-1}}^b = \alpha_j^c \ldots \alpha_{j_{h-1}}^d \ (\neq \varepsilon), \tag{19}$$

wobei jeder Exponent entweder durch $p$ nicht teilbar oder gleich 0 ist. Wählt man also die $t_i$ im Hilfssatz 3 passend, so treten beide Seiten von (19) beim Ausmultiplizieren der rechten Seite von (15) auf. Da aber diese (mit (17) zusammen) ein direktes Produkt ist und $\alpha_i, \ldots, \alpha_{i_{g-1}}, \alpha_j, \ldots, \alpha_{j_{h-1}}$ verschieden sind, so ist (19) unmöglich. Dieser Widerspruch beweist Hilfssatz 5.

Nunmehr können wir Satz 2 auch im Fall beweisen, daß $G$ keine $p$-Gruppe ist. Zu diesem Zweck machen wir die Induktionsannahme, daß Satz 2 für die Gruppen von kleinerer Ordnung als $G$ richtig ist, ferner nehmen wir an, daß Satz 1 für eine prime Hajóssche Zerlegung (17) von $G$ falsch ist. Nach der Reduktion am Ende des § 2 genügt es hieraus einen Widerspruch abzuleiten.

Aus Hilfssatz 3 folgt sofort, daß (mit veränderten Bezeichnungen) auch eine (prime) Hajóssche Zerlegung

$$G = (\alpha_1\beta_1)_{p_1} \ldots (\alpha_m\beta_m)_{p_m} (\gamma_1)_{r_1} \ldots (\gamma_n)_{r_n} \qquad (m, n \geqq 0;\ m+n > 0), \tag{20}$$

gilt, wobei

$$o(\alpha_i) = p_i \neq o(\beta_i) = q_i \qquad (i = 1, \ldots, m), \tag{21}$$

(22) $\quad o(\gamma_i)$ eine durch $r_i^2$ teilbare Potenz von $r_i$ ist $(i = 1, \ldots, n)$.

Mit (20) zusammen sind nach Hilfssatz 3 auch

$$G = (\alpha_1^{a_1}\beta_1^{b_1})_{p_1} \ldots (\alpha_m^{a_m}\beta_m^{b_m})_{p_m} (\gamma_1^{c_1})_{r_1} \ldots (\gamma_n^{c_n})_{r_n} \tag{23}$$
$$(p_i \nmid a_i,\ r_k \nmid c_k;\ i = 1, \ldots, m;\ k = 1, \ldots, n)$$

lauter Hajóssche Zerlegungen.

Wenn in (20) $m=0$ wäre, so liefert ein geeignetes Teilprodukt von (20) eine Hajóssche Zerlegung einer Sylowgruppe von $G$. Das ist aber wegen (22) unmöglich, da doch Satz 2 und noch mehr Satz 1 für $p$-Gruppen richtig ist. Folglich muß $m \geqq 1$ gelten.

Wir zeigen, daß es zu jedem $i(=1,\ldots,m)$ mindestens ein geordnetes Normalsystem

$$(24) \qquad \begin{pmatrix} \alpha_i & \gamma_s & \gamma_t & \ldots & \gamma_w \\ p_i & p_i & p_i & \ldots & p_i \end{pmatrix} \qquad (1 \leqq s, t, \ldots, x \leqq n)$$

und ein von $i$ verschiedenes $i'(=1,\ldots,m)$ mit

$$(25) \qquad \beta_{i'} \in \{\alpha_i, \gamma_s, \gamma_t, \ldots, \gamma_w\} \qquad (i' \neq i)$$

gibt.

Als einfacher Spezialfall von (23) gilt nämlich (20) auch nach Ersetzung des $i$-ten Faktors durch $(\alpha_i)_{p_i}$. Da dieser nach (21) eine Gruppe ist, so ist für die entstandene Hajóssche Zerlegung von $G$ Satz 1, also wegen der Induktionsannahme und Hilfssatz 1 auch Satz 2 richtig. Hiernach ist das dieser Zerlegung entsprechende System (1) normal. Letzteres geht nach passender Spaltenvertauschung in ein geordnetes Normalsystem über, das offenbar von der Form

$$(26) \qquad \begin{pmatrix} \alpha_i & \gamma_s & \gamma_t & \ldots & \gamma_w \alpha_{i'} & \beta_{i'} & \ldots \\ p_i & p_i & p_i & \ldots & p_i & p_{i'} & \ldots \end{pmatrix} \qquad (i' \neq i)$$

sein muß (die $\gamma_s, \gamma_t, \ldots, \gamma_w$ können auch fehlen), wobei sich mit Notwendigkeit auch $m \geqq 2$ ergab. Hieraus folgt, daß auch (24) ein geordnetes Normalsystem und

$$(\alpha_{i'}\beta_{i'})^{p_{i'}} \in \{\alpha_i, \gamma_s, \gamma_t, \ldots, \gamma_x\}$$

ist. Hier fällt $\alpha_{i'}$ wegen (21) heraus. Ferner ist nach (24) der Exponent $p_{i'}$ prim zu $o(\beta_{i'})$, weshalb in der Tat auch (25) richtig ist.

Nun denken wir uns zu jedem $i(=1,\ldots,m)$ ein geordnetes Normalsystem (24) und ein $i'(=1,\ldots,m)$ mit der Eigenschaft (25) festgewählt. Insbesondere ist dann

$$i \to i' \qquad (i=1,\ldots,m)$$

eine Abbildung der Menge der Zahlen $1,\ldots,m$ in sich ohne Fixelemente. Folglich gibt es eine Untermenge dieser Menge mit mindestens zwei Elementen, die dabei permutiert wird. Diese Permutation läßt sich nach passender Umordnung der Faktoren von (20) durch

$$(27) \qquad i \to i' \qquad (i=1,\ldots,m_0;\ 2 \leqq m_0 \leqq m)$$

angeben.

Für jedes $i(=1,\ldots,m_0)$ bezeichnen wir mit $M_i$ die Menge der Elemente in der ersten Zeile von (24). Da nach (23) insbesondere

$$G = (\alpha_1)_{p_1} \ldots (\alpha_m)_{p_m} (\gamma_1)_{r_1} \ldots (\gamma_n)_{r_n}$$

gilt, so folgt aus Hilfssatz 5, daß die $M_i (i=1, \ldots, m_0)$ paarweise elementfremd sind.

Wird andererseits Hilfssatz 4 auf das geordnete Normalsystem (24) und das Element $\beta_{i'}$ in (25) angewendet, so entsteht eine Gleichung von der Form

$$\beta_{i'} = \alpha_i^{a_i} \mu_1^{c_1} \ldots \mu_k^{c_k} \qquad (i=1, \ldots, m_0),$$

wobei $\mu_1, \ldots, \mu_k (\neq \alpha_i)$ verschiedene Elemente von $M_i$ sind und jeder Exponent prim zur Ordnung der Basis ist. Nach Multiplizieren dieser Gleichungen entsteht wegen (27) und des vorher Gesagten eine Gleichung

$$(\alpha_1^{a_1} \beta_1^{-1}) \ldots (\alpha_{m_0}^{a_{m_0}} \beta_{m_0}^{-1}) \gamma_1^{d_1} \ldots \gamma_n^{d_n} = \varepsilon,$$

wobei jedes $d_k$ gleich 0 oder zu $o(\gamma_k)$ prim ist. Diese Gleichung steht aber im Widerspruch damit, daß (23) ein direktes Produkt ist. Das beendet den Beweis des Satzes von HAJÓS in seiner im Satz 2 gefaßten scharfen Form.

## § 5. Schlußbemerkungen

Wir wollen kurz skizzieren, wie sich der einzige ringtheoretische Schluß im obigen Beweise des Satzes von HAJÓS der Bemerkung von Herrn SZELE folgend gruppentheoretisch ausführen läßt, wodurch der ganze Beweis einen rein gruppentheoretischen Charakter erhält. Hierzu haben wir folgende Behauptung auf gruppentheoretischem Wege zu beweisen: Die Annahme, daß (12), (14) Hajóssche Zerlegungen von $G$ sind und (13) gilt, ist widersprüchig.

Nach (12) gilt nämlich allgemeiner

$$G = \prod_{i=1}^{m} \alpha_i^{p a_i + x_i} (\alpha_i)_p \prod_{k=m+1}^{n} (\alpha_k)_p \tag{28}$$

$$\left(a_i = 0, \ldots, \frac{s_i}{p} - 1;\; x_i = 0, 1;\; i = 1, \ldots, m\right).$$

(Das ist sogar für beliebige ganze Zahlen $a_i, x_i$ richtig.) Das Produkt auf der rechten Seite bezeichnen wir mit $P(x) = P(x_1, \ldots, x_m)$ und nennen es gerade oder ungerade, je nachdem $x_1 + \cdots + x_m$ gerade oder ungerade ist. (In der Wahrheit hängt $P(x)$ auch von $a_1, \ldots, a_m$ ab.) Es ist klar, daß (nach Ausmultiplizieren) die Elemente von $P(x)$ von der Form sind:

$$\alpha_1^{y_1} \ldots \alpha_m^{y_m} \alpha_{m+1}^{z_{m+1}} \ldots \alpha_n^{z_n} \qquad \begin{pmatrix} 0 \leqq y_i \leqq s_i; & i = 1, \ldots, m \\ 0 \leqq z_k \leqq p-1; & k = m+1, \ldots, n \end{pmatrix}. \tag{29}$$

Ist jedes $y_i$ gleich 0 oder $s_i$, so gehört (29) offenbar nur einem $P(x)$ zu, und zwar ist

$$\alpha_1^{s_1 x_1} \ldots \alpha_m^{s_m x_m} \alpha_{m+1}^{z_{m+1}} \ldots \alpha_n^{z_n} \tag{30}$$

in einem $P(x)$ enthalten. Wegen (14) sind diese Elemente (30) auch verschieden. Wir teilen sie in zwei Komplexe $A_0, A_1$ ein, je nachdem $x_1 + \cdots + x_m$

gerade oder ungerade ist. Wegen des Gesagten ist jedenfalls $A_0 \neq A_1$. Der gewünschte Widerspruch wird so entstehen, daß wir $A_0 = A_1$ zeigen. Wir betrachten hierzu auch die übrigen Elemente (29), in denen nämlich mindestens ein $y_i$ mit $0 < y_i < s_i$ vorkommt. Man sieht leicht, daß jedes solche Element gleichvielen geraden und ungeraden $P(x)$ zugehört. Da die Anzahl der geraden und ungeraden $P(x)$ gleich ist und nach (28) alle $P(x)$ gleich $G$ sind, so folgt $A_0 = A_1$ in der Tat, womit der Beweis beendet ist. (Freilich war obiger ringtheoretischer Beweis viel eleganter.)

Unser Lemma im § 3 verdient auch einiges selbständiges Interesse. Seine unmittelbare Verallgemeinerung würde so lauten: Wenn

$$O\{\alpha_1, \ldots, \alpha_n\} = p_1 \ldots p_n,$$
$$p_{i_1} \ldots p_{i_k} \mid O\{\alpha_{i_1}, \ldots, \alpha_{i_k}\} \qquad (1 \leqq i_1 < \cdots < i_k \leqq n;\ 1 \leqq k \leqq n-1)$$

gelten, so gibt es ganze Zahlen $s_1, \ldots, s_n$, für die

$$\begin{pmatrix} \alpha_1^{s_1} & \ldots & \alpha^{s_n} \\ p_1 & \ldots & p_n \end{pmatrix}$$

ein Normalsystem ist. Wir wissen nicht, ob diese Verallgemeinerung richtig ist. Wenn das auch der Fall ist, so könnte man dadurch unseren Beweis (für nicht-$p$-Gruppen) doch nicht verkürzen. Insofern ist unser Lemma „scharf", als seine weitere Verallgemeinerung auf den Fall beliebiger ganzen Zahlen $e_1, \ldots, e_n (\geqq 2)$ statt $p_1, \ldots, p_n$ sogar schon dann falsch ist, wenn dabei $e_1, \ldots, e_n$ Potenzen von $p$ sind, was wir mit folgendem Beispiel zeigen: Es bestehe $o(\varrho) = p^3$, $o(\sigma) = p$, $\sigma \notin \{\varrho\}$. Für $\alpha_1 = \varrho$, $\alpha_2 = \varrho\sigma$ gelten dann

$$O\{\alpha_1, \alpha_2\} = p^2 \cdot p^2,\ p^2 \mid O\{\alpha_1\}, O\{\alpha_2\},$$

trotzdem ist

$$\begin{pmatrix} \alpha_1^{s_1} & \alpha_2^{s_2} \\ p^2 & p^2 \end{pmatrix}$$

für keine ganzen Zahlen $s_1, s_2$ ein Normalsystem. Hieraus folgt, daß in unserem Beweis die im Hilfssatz 2 gefaßte Reduktion auf den Fall von Primsimplexen unterbehrlich ist.

*(Eingegangen am 29. Januar 1955.)*

## Literaturverzeichnis

[1] G. Hajós, Többméretű terek egyszeres befedése kockaráccsal, *Matematikai és fizikai lapok*, 48 (1941), S. 37—64 (ungarisch mit deutschem Auszug).

[2] G. Hajós, Über einfache und mehrfache Bedeckung des $n$-dimensionalen Raumes mit einem Würfelgitter, *Math. Zeitschrift*, 47 (1942), S. 427—467.

[3] L. Rédei, Jelentés az 1942. évi Kőnig Gyula jutalomról: Hajós György munkáinak ismertetése, *Matematikai és fizikai lapok*, **49** (1942), S. 1—16. (Ein ungarisch verfaßter Bericht über die Arbeiten von G. Hajós.)
[4] L. Rédei, Vereinfachter Beweis des Satzes von Minkowski—Hajós, *Acta Sci. Math.*, **13** (1949), S. 21—35.
[5] L. Rédei, Kurzer Beweis des gruppentheoretischen Satzes von Hajós, *Comm. Math. Helv.*, **23** (1949), S. 272—282.
[6] L. Rédei, Die Reduktion des gruppentheoretischen Satzes von Hajós auf den Fall von $p$-Gruppen, *Monatshefte für Math.*, **53** (1949), S. 221—226.
[7] L. Rédei, *Algebra I* (Budapest, 1954) (ungarisch).
[8] L. Rédei, Die gruppentheoretischen Zetafunktionen und der Satz von Hajós, *Acta Math. Acad. Sci. Hung.*, **6** (1955) (im Erscheinen).
[9] T. Szele, Neuer Vereinfachter Beweis des gruppentheoretischen Satzes von Hajós, *Publicationes Math. (Debrecen)*, **1** (1949), S. 56—62.

## НОВОЕ ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ ХАЙОША О КОНЕЧНЫХ АБЕЛЕВЫХ ГРУППАХ

Л. Редеи (Сегед)

(Резюме)

Пусть $G(\neq\varepsilon)$ — конечная абелева группа, $\varepsilon$ же единица $G$. Обозначим порядок $G$ через $O(G)$ а порядок элемента $\alpha(\in G)$ через $o(\alpha)$. Если $o(\alpha)\geqq e\geqq 2$, то обозначим множество элементов $\varepsilon, \alpha, \ldots, \alpha^{e-1}$ через $(\alpha)_e$. Оно образует группу очевидно тогда и только тогда, когда $o(\alpha)=e$. Знаменимая теорема Хайоша, доказывающая гипотезу Минковского „о предельном случае" системы линейных неравенств, утверждает, что

(1) $$G=(\alpha_1)_{e_1}\cdots(\alpha_k)_{e_k},\quad O(G)=e_1\ldots e_k$$

возможно лишь если хотя бы одно из множеств $(\alpha_i)_{e_i}$ образует группу. Для этой теоремы имелось до сих пор только оригинальное доказательство Хайоша (см. Хайош [2]), упрощенное Редеи [5] и Селе [9]. Излагаемое ниже простое доказательство использует прежде всего простой факт, также принадлежащий Хайошу, что теорему достаточно доказать для частного случая

(2) $$G=(\alpha_1)_{p_1}\cdots(\alpha_1)_{p_n},\quad O(G)=p_1\cdots p_n,$$

где $p_1,\ldots,p_n$ простые числа.

Дальнейшим шагом является следующая нетрудная

Лемма. Если $\alpha_1,\ldots,\alpha_n\in G$, $p$ простое число и

(3) $$O(\{\alpha_1,\ldots,\alpha_n\})=p,$$

(4) $$O(\{\alpha_{i_1},\ldots,\alpha_{i_k}\})\geqq p^k\qquad(1\leqq i_1<\cdots<i_k\leqq n;\ 1\leqq k\leqq n-1),$$

то, удобно располагая элементы $\alpha_1,\ldots,\alpha_n$, можно указать такие степени $s_1,\ldots,s_n(\geqq 1)$ числа $p$, что

(5) $$O(\{\alpha_1^{s_1},\ldots,\alpha_k^{s_k}\})=p^k\qquad(1\leqq k\leqq n).$$

($\{\alpha,\beta,\ldots\}$ означает группу, порожденную элементами $\alpha,\beta,\ldots$.)

Если, в частности, $G$ является $p$-группой, то в соответствии с вышесказанным можно ограничиться исследованием случая

(6) $$G=(\alpha_1)_p\cdots(\alpha_n)_p,\quad O(G)=p^n.$$

Очевидно, что из (6) следуют соотношения (3) и (4). Так как далее (6) справедливо при любом расположении факторов, то можно предположить и (5) выполненным. Если теперь $s_1=1$, то из частного случая (5) соответствующего $k=1$ уже вытекает $o(\alpha_1)=p$ так, что $(\alpha_1)_p$ образует группу. В оставшемся случае можно считать (при удобно выбранном $m$), что

(7) $$p\,|\,s_1, \ldots, s_m;\quad s_{m+1}=\cdots=s_n=1 \qquad (1\leqq m\leqq n).$$

С другой стороны из (5) сразу следует

(8) $$G=(\alpha_1^{s_1})_p\ldots(\alpha_n^{s_n})_p.$$

Однако мы из соотношений (6)—(8) выведем противоречие.

Для этой цели используем групповое кольцо $\mathfrak{J}(G)$ группы $G$, где $\mathfrak{J}$ означает кольцо целых чисел. Обозначим через $[\alpha]_r$ сумму $\varepsilon+\alpha+\cdots+\alpha^{r-1}\,(\in\mathfrak{J}(G))$. Так как $\alpha_1 G=G$ и $(\varepsilon-\alpha_1)[\alpha_1]_p=\varepsilon-\alpha_1^p$, то из (6) сразу следует

$$(\varepsilon-\alpha_1^p)[\alpha_2]_p\cdots[\alpha_n]_p=0,$$

ввиду (7) тем более

$$(\varepsilon-\alpha_1^{s_1})\cdots(\varepsilon-\alpha_m^{s_m})[\alpha_{m+1}]_p\cdots[\alpha_n]_p=0.$$

На левой стороне равенства появится (после умножения) член $\varepsilon$, которой и равен, с точностью до знака, дальнейшему члену. Ввиду (7) это действительно противоречит (8).

Случай групп не являющихся $p$-группами можно трактовать опираясь на этот случай (см. Редеи [6]). К сожалению эта часть доказательства, только принципиально проста, подробное доказательство довольно сложно.

# SUR LA CARACTÉRISATION DE CERTAINES CLASSES DE FONCTIONS AU SENS DE LA THÉORIE CONSTRUCTIVE DES FONCTIONS

Par

G. ALEXITS (Budapest), membre de l'Académie

Un problème très attirant de la théorie constructive des fonctions consiste dans la caractérisation de la structure des fonctions par l'ordre de grandeur qu'on obtient en les approchant avec les moyennes $(C, \alpha)$ de leurs séries de Fourier. Il y a quelques années nous avons réussi à caractériser de cette manière[1] les classes[2] Lip $(1, p)$ en étudiant l'approximation dans l'espace $L^p$. Or l'approximation dans l'espace $L^p$ ne donne aucun renseignement sur la vitesse de la convergence atteinte aux différents points de l'intervalle $(0, 2\pi)$. Toutefois, une caractérisation de ces classes par l'ordre de grandeur de l'approximation dans le sens habituel aurait un intérêt, parce qu'elle éclaircirait en détail les propriétés structurelles des fonctions appartenantes à ces classes. De plus, la méthode qui pourrait servir à une telle caractérisation permetterait, en même temps, d'étudier aussi l'ordre de grandeur de l'approximation des fonctions qui ne sont dérivables qu'en certains points. Une étude méthodique de l'approximation aux points individuels pourrait conduire, peut-être, à des résultats plus profonds que ceux qu'on obtient à l'aide des recherches concernant la vitesse de l'approximation en moyenne dans l'intervalle entier $(0, 2\pi)$.

Dans les suivants, nous allons résoudre le problème d'une telle caractérisation des classes Lip $(1, p)$ avec $p > 1$ et essayons de faire les premiers pas vers une étude systématique de l'ordre de grandeur obtenu point par

[1] G. Alexits, Sur l'ordre de grandeur de l'approximation d'une fonction périodique par les sommes de Fejér, *Acta Math. Acad. Sci. Hung.*, 3 (1952), p. 29—40 (en français) et p. 41—42 (en russe).

[2] Une fonction $f(x)$ appartient à la classe Lip $(\alpha, p)$, si

$$\sup_{|h| \leq \delta} \left\{ \int_0^{2\pi} |f(x+h) - f(x)|^p dx \right\}^{\frac{1}{p}} = O(\delta^\alpha)$$

où $0 < \alpha \leq 1$ et $p \geqq 1$.

point en approximant des fonctions qui ne sont dérivables qu'en certains points. La méthode que nous allons employer consiste dans l'application d'un lemme que nous avons démontré ailleurs[3] et qui nous a servi à nos recherches plus ou moins semblables.

## 1. Caractérisation des classes Lip $(1, p)$

Désignons par

$$\sigma_n^\alpha(x) = \frac{a_0}{2} + \sum_{k=1}^{n} \frac{\binom{n-k+\alpha}{n-k}}{\binom{n+\alpha}{n}} (a_k \cos kx + b_k \sin kx)$$

la $n$-ième moyenne $(C, \alpha)$ de la série de Fourier d'une fonction intégrable $f(x)$. La fonction conjuguée sera désignée par $\tilde{f}(x)$ et

$$\tilde{\sigma}_n^\alpha(x) = \sum_{k=1}^{n} \frac{\binom{n-k+\alpha}{n-k}}{\binom{n+\alpha}{n}} (b_k \cos kx - a_k \sin kx)$$

est la $n$-iéme moyenne $(C, \alpha)$ de la série conjuguée. Le lemme dont nous avons fait mention et que nous avons démontré ailleurs est le suivant:

*Si, pour un* $\alpha > 0$, *on a en un point* $x$

$$|\tilde{f}(x) - \tilde{\sigma}_n^\alpha(x)| \leqq \frac{K}{n},$$

*alors il existe une constante* $C_\alpha$ *ne dépendant que de* $\alpha$ *telle que* $|\sigma_n^{\alpha\prime}(x)| \leqq C_\alpha K$. *Inversement, si les dérivées* $\sigma_n^{\alpha\prime}(x)$ *des moyennes* $\sigma_n^\alpha(x)$ *satisfont au point* $x$ *à la condition* $|\sigma_n^{\alpha\prime}(x)| \leqq K$, *alors*

$$|\tilde{f}(x) - \tilde{\sigma}_n^\alpha(x)| \leqq \frac{C_\alpha K}{n}.$$

A l'aide de ce lemme, nous pouvons obtenir une caractérisation complète des classes Lip $(1, p)$ de la manière suivante:

Théorème 1. *Pour que la fonction* $f(x)$ *appartienne à la classe* Lip $(1, p)$ *où* $p \leqq 1$, *il faut qu'il existe, pour tout* $\alpha > 0$, *une fonction* $F_\alpha(x) \in L^p$ *telle que*

$$|f(x) - \sigma_n^\alpha(x)| \leqq \frac{F_\alpha(x)}{n}, \quad |\tilde{f}(x) - \tilde{\sigma}_n^\alpha(x)| \leqq \frac{F_\alpha(x)}{n},$$

*et il suffit qu'une au moins de ces inégalités soit satisfaite pour un* $\alpha > 0$ *arbitrairement fixé.*

[3] L. c. [1] et G. Alexits, A Fourier-sor Cesàro-közepeivel való approximáció nagyságrendjéről, *Mat. Fiz. Lapok*, **48** (1941), p. 410—421 (hongrois) et p. 421—422 (français).

Soit, en effet, $f \in \mathrm{Lip}\,(1, p)$. Il existe alors une fonction absolument continue $g(x)$ qui ne diffère de $f(x)$ qu'aux points d'un ensemble de mesure nulle et dont la dérivée $g'(x)$ est $L^p$-intégrable.[4] Alors $\sigma_n^{\alpha\prime}(x)$ est la $n$-ième moyenne $(C, \alpha)$ de la série de Fourier de $g'(x)$; il existe[5] donc pour tout $\alpha > 0$ une fonction $L^p$-intégrable $\Phi_\alpha(x)$ telle que

$$|\sigma_n^{\alpha\prime}(x)| \leqq \Phi_\alpha(x) \qquad (n = 1, 2, \ldots).$$

En appliquant notre lemme, il s'ensuit que

$$|\tilde{g}(x) - \tilde{\sigma}_n^\alpha(x)| \leqq \frac{C_\alpha \Phi_\alpha(x)}{n}.$$

Mais $\tilde{g}(x)$ ne diffère de $\tilde{f}(x)$ qu'en un ensemble de mesure nulle; en changeant donc les valeurs de $\Phi_\alpha(x)$ aux points de cet ensemble, on obtient aussi

$$|\tilde{f}(x) - \tilde{\sigma}_n^\alpha(x)| \leqq \frac{C_\alpha \Phi_\alpha(x)}{n}.$$

C'est la deuxième des inégalités que nous voulons démontrer. Quant à la première, on n'a qu'à observer que $f(x)$ et $\tilde{f}(x)$ appartiennent en même temps à la classe $\mathrm{Lip}\,(1, p)$. En effet, puisque $g' \in L^p$, il s'ensuit[6] pour sa conjuguée: $\tilde{g}' \in L^p$. Or cette relation entraîne[4] que l'intégrale de $\tilde{g}'(x)$, c'est à dire que $\tilde{g}(x)$ appartient à $\mathrm{Lip}\,(1, p)$. Mais $\tilde{g}(x)$ ne diffère qu'en un ensemble de mesure nulle de $\tilde{f}(x)$, par conséquent $\tilde{f} \in \mathrm{Lip}\,(1, p)$. On en obtient pour la conjuguée $f(x)$ de la fonction $\tilde{f}(x)$ comme auparavant

$$|f(x) - \sigma_n^\alpha(x)| \leqq \frac{C_\alpha \Psi_\alpha(x)}{n}$$

où $\Psi_\alpha(x)$ est $L^p$-intégrable. En posant $F_\alpha(x) = \max\{C_\alpha \Phi_\alpha(x), C_\alpha \Psi_\alpha(x)\}$, la nécessité de notre condition est démontrée. — Quant à sa suffisance, soit une de nos inégalités, par exemple la première, satisfaite. On en tire par l'application de notre lemme

$$|\tilde{\sigma}_n^{\alpha\prime}(x)| \leqq C_\alpha F_\alpha(x).$$

Les sommes $\tilde{\sigma}_n^{\alpha\prime}(x)$ sont donc, d'après des théorèmes connus,[7] les moyennes $(C, \alpha)$ de la série de Fourier d'une fonction $L^p$-intégrable et de même les sommes $\sigma_n^{\alpha\prime}(x)$ conjuguées aux $\tilde{\sigma}_n^{\alpha\prime}(x)$. Les moyennes $\sigma_n^{\alpha\prime}(x)$ convergent donc presque partout vers une fonction $L^p$-intégrable $g'(x)$ dont l'intégrale $g(x)$, ayant presque partout une dérivée $L^p$-intégrable, appartient à la classe

[4] G. H. Hardy—J. E. Littlewood, Some properties of fractional integrals. I, *Math. Zeitschrift*, **27** (1928), p. 565—606.

[5] G. H. Hardy—J. E. Littlewood, A maximal theorem with function-theoretic applications, *Acta Math.*, **54** (1930), p. 81—116.

[6] M. Riesz, Sur les fonctions conjuguées, *Math. Zeitschrift*, **27** (1928), p. 218—249.

[7] V. p. ex. A. Zygmund, *Trigonometrical series* (Warszawa—Lwów, 1935), p. 87.

Lip$(1,p)$. Mais $f(x)$ et $g(x)$ ont toutes les deux $\sigma_n^\alpha(x)$ pour $n$-ième moyenne $(C,\alpha)$ de leur série de Fourier, elles ne diffèrent donc qu'en un ensemble de mesure nulle, par conséquent $f\in$ Lip$(1,p)$ et notre théorème est entièrement démontré.

Ce théorème n'est plus valable pour $p=1$, parce qu'alors aucune de nos inégalités n'est nécessaire et ce n'est que la deuxième inégalité qui assure l'appartenance de $f(x)$ à la classe Lip$(1,1)$.

## 2. L'ordre de grandeur de l'approximation des fonctions conjuguées à des fonctions dérivables en certains points

Les précédents contiennent, au fond, la caractérisation des fonctions dérivables presque partout dont les dérivées sont soumises à des conditions d'intégrabilité. En renonçant à ces conditions, nous ne pouvons plus caractériser les classes plus larges, mais la dérivabilité seule, même si elle n'est satisfaite qu'en un seul point, entraîne déjà une convergence rapide des moyennes $\sigma_n^\alpha(x)$ à ce point.

Théorème 2. *Si la fonction $f(x)$ est dérivable au point $x_0$, alors on a pour tout $\alpha>1$*

$$\tilde{f}(x_0)-\tilde{\sigma}_n^\alpha(x_0)=O\left(\frac{1}{n}\right).$$

En effet, puisque la dérivée $f'(x_0)$ existe, les dérivées $\sigma_n^{\alpha\prime}(x_0)$ convergent[8] vers $f'(x_0)$, par conséquent

$$\sigma_n^{\alpha\prime}(x_0)=O(1)$$

et la proposition résulte immédiatement de notre lemme.

On voit que $f(x)$ étant dérivable aux points d'un ensemble, le degré d'approximation obtenu à l'aide des moyennes $\tilde{\sigma}_n^\alpha(x)$ peut être évalué par $F(x)/n$ où $F(x)$ est une fonction finie sur cet ensemble, mais nous ne savons rien sur la sturcture de $F(x)$. On obtient un peu plus en supposant $f\in$ Lip $(1,p)$ dans un intervalle $(a,b)\subset(0,2\pi)$.

Théorème 3. *Si $f\in$ Lip$(1,p)$ dans un intervalle $(a,b)$ pour un $p>1$, alors il existe, pour tout $\varepsilon>0$, une fonction $F_\varepsilon(x)\in L^p$, telle que*

$$|\tilde{f}(x)-\tilde{\sigma}_n^1(x)|\leqq\frac{F_\varepsilon(x)}{n}$$

*aux points $x$ de l'intervalle $(a+\varepsilon, b-\varepsilon)$.*

Comme $f\in$ Lip$(1,p)$ dans l'intervalle $(a,b)$, il existe[4] une fonction $g(x)$ équivalente à $f(x)$ en $(a,b)$ et telle que sa dérivée $g'(x)$ soit $\in L^p(a,b)$.

[8] V. p. ex. l. c. [7], p. 55.

Posons $\bar{g}(x)=g(x)$ pour $x\in(a,b)$ et $\bar{g}(x)=0$ ailleurs. Alors $\bar{g}'(x)$ existe presque partout et $\bar{g}'\in L^p$. En désignant donc par $\sigma_n(\bar{g},x)$ la $n$-ième moyenne de FEJÉR de la série de Fourier de $\bar{g}(x)$, on obtient comme auparavant

$$|\sigma_n'(\bar{g},x)|\leqq G(x)$$

où $G\in L^p$. Mais, en désignant par $K_n(t,x)$ le noyau de FEJÉR, on obtient aussi

$$|\sigma_n^{1\prime}(x)-\sigma_n'(\bar{g},x)|\leqq\int_0^{2\pi}|f(t)-\bar{g}(t)||K_n'(t,x)|\,dt=$$

$$=\left(\int_0^a+\int_b^{2\pi}\right)|f(t)-\bar{g}(t)|\,O\left(\frac{1}{(t-x)^3}\right)dt.$$

On a donc pour tout $\varepsilon>0$ et pour tous les points $x\in(a+\varepsilon,b-\varepsilon)$

$$|\sigma_n^{1\prime}(x)-\sigma_n'(\bar{g},x)|=O\left(\frac{1}{\varepsilon^3}\right)\int_0^{2\pi}|f(t)-\bar{g}(t)|\,dt=K_\varepsilon.$$

Par conséquent

$$|\sigma_n^{1\prime}(x)|\leqq K_\varepsilon+|\sigma_n'(\bar{g},x)|\leqq K_\varepsilon+G(x)=F_\varepsilon(x)$$

où la fonction $F_\varepsilon(x)$ est évidemment $L^p$-intégrable. Il s'ensuit par l'application de notre lemme:

$$|\tilde{f}(x)-\tilde{\sigma}_n^1(x)|\leqq\frac{F_\varepsilon(x)}{n}, \qquad \text{c. q. f. d.}$$

*(Reçu le 7 février 1955.)*

## ОХАРАКТЕРИЗОВАНИЕ НЕКОТОРЫХ КЛАССОВ ФУНКЦИЙ С ПОМОЩЬЮ КОНСТРУКТИВНОЙ ТЕОРИИ ФУНКЦИЙ

Г. Алексич (Будапешт)

(Резюме)

Функции принадлежащие классу Lip $(1,p)$ были охарактеризованы авторам быстротой сходимости в пространстве $L^p$ средних Фейера их сопряженных рядов Фурье. В настоящей работе автор занимается охарактеризованием класса Lip $(1,p)$ с помощью порядка величины почечной аппроксимации. Обозначим для этой цели через $\sigma_n^\alpha(x)$ средние $(C,\alpha)$ порядка $\alpha>0$ ряда Фурье интегрируемой функции $f(x)$, а через $\tilde{f}(x)$ и $\tilde{\sigma}_n^\alpha(x)$ сопряженные фуикций $f(x)$ и $\sigma_n^\alpha(x)$.

I. Необходимым для того, чтобы $f(x)$ принадлежала классу Lip $(1,p)$ $(p>1)$, является существование для всех $\alpha>0$ такой функции $F_\alpha(x)$ интегрируемой $L^p$, для которой

$$|f(x)-\sigma_n^\alpha(x)|\leqq\frac{F_\alpha(x)}{n},\quad |\tilde{f}(x)-\tilde{\sigma}_n^\alpha(x)|\leqq\frac{F_\alpha(x)}{n};$$

выполнение хотя бы одного из этих неравенств является достаточным для того, чтобы было $f \in \mathrm{Lip}\,(1, p)$.

Соотношение $f \in \mathrm{Lip}\,(1, p)$ по существу означает, что $f(x)$ почти всюду имеет производную, интегрируемую $L^p$. Если отбросим требование интегрируемости, то можем оценить лишь быстроту сходимости аппроксимации.

II. Если $f(x)$ имеет производную в точке $x_0$, то для всех $\alpha > 1$

$$\tilde{f}(x_0) - \tilde{\sigma}_n^{\alpha}(x_0) = O\left(\frac{1}{n}\right).$$

Если $f \in \mathrm{Lip}\,(1, p)$ в некотором подинтервале интервала $(0, 2\pi)$, то имеет место следующее более сильное утверждение:

III. Если $f \in \mathrm{Lip}\,(1, p)$ в интервале $(a, b) \subset (0, 2\pi)$ $(p > 1)$, то при любом $\varepsilon > 0$ можно задать функцию $F_\varepsilon(x)$ интегрируемую $L^p$, для которой

$$|\tilde{f}(x) - \tilde{\sigma}_n^1(x)| \leqq \frac{F_\varepsilon(x)}{n}$$

во всех точках интервала $(a + \varepsilon, b - \varepsilon)$.

# ON THE ROLE OF THE LEBESGUE FUNCTIONS IN THE THEORY OF THE LAGRANGE INTERPOLATION [1]

By

P. ERDŐS (Jerusalem) and P. TURÁN (Budapest), member of the Academy

*To Prof. dr. L. Fejér on the occasion of his 75th birthday*

**1.** Let there be given a triangular matrix

$$A \equiv \begin{array}{llll} x_{11} & & & \\ x_{12} & x_{22} & & \\ \vdots & & \ddots & \\ x_{1n} & x_{2n} & \dots & x_{nn} \\ \vdots & & & \ddots \end{array} \tag{1.1}$$

where for $n=1,2,\dots$ we have

$$1 \geqq x_{1n} > x_{2n} > \dots > x_{nn} \geqq -1. \tag{1.2}$$

Then, as it is well known, for given values $y_{\nu n}$ there is exactly one polynomial $g(x)$ of degree $\leqq n-1$ such that

$$g(x_{\nu n}) = y_{\nu n} \qquad (\nu = 1, 2, \dots, n).$$

If the values $y_{\nu n}$ are the values $f(x_{\nu n})$ of a function $f(x)$ defined in $[-1, +1]$, then we call the corresponding $g(x)$ polynomial "the $n^{th}$ interpolatory polynomial of $f(x)$ belonging to $A$" and denote it by $L_n(f, A)$ or — if misunderstanding cannot arise — by $L_n(f)$. The abscissae $x_{\nu n}$ are called the $n^{th}$ fundamental points of the matrix $A$ and are sometimes denoted also by $x_\nu$. It is well known that $L_n(f, A)$ can be written in the form

$$L_n(f, A) = \sum_{\nu=1}^{n} f(x_{\nu n})\, l_{\nu n}(x, A), \tag{1.3}$$

where the polynomials $l_{\nu n}(x, A)$, the so-called fundamental-functions belonging

[1] A part of the results (assertions a) and b) of this paper) was the subject of a lecture made by one of us at a colloquium for the constructive function-theory in Eger (Hungary), 29 Nov. 1953; they were found twenty years ago. The new results showing they are best possibles were a subject of another lecture in Pécs (Hungary), 18 Sept. 1954.

to $A$, depend only upon $A$ and have the representation

$$l_{\nu n}(x, A) = \frac{\omega_n(x, A)}{\omega_n'(x_{\nu n}, A)\ (x - x_{\nu n})} \tag{1.4}$$

where

$$\omega_n(x, A) = \prod_{\nu=1}^{n} (x - x_{\nu n}). \tag{1.5}$$

As easy to verify, we have

$$\sum_{\nu=1}^{n} l_{\nu n}(x, A) \equiv 1 \tag{1.6}$$

and if $h(x)$ denotes an arbitrary polynomial of degree $\leqq n-1$, then

$$L_n(h, A) \equiv h(x). \tag{1.7}$$

From (1.3) it follows that for an $f(x)$ bounded in $[-1, +1]$ we have for $-1 \leqq x \leqq +1$

$$|L_n(f, A)| \leqq \left(\sum_{\nu=1}^{n} |l_{\nu n}(x, A)|\right) \sup_{-1 \leqq x \leqq +1} |f(x)|. \tag{1.8}$$

2. One would be inclined to think that if $A$ is such that for an arbitrarily small $\varepsilon > 0$

$$x_{\nu n} - x_{\nu+1, n} \leqq \varepsilon, \quad x_{0n} = -x_{n+1, n} = 1 \qquad (\nu = 0, 1, \ldots, n;\ n > n_0(\varepsilon)), \tag{2.1}$$

then the sequence $L_n(f, A)$ converges uniformly in $[-1, +1]$ to $f(x)$ whenever[2] $f(x) \in C$. It was a great surprise at the end of the last century when RUNGE and BOREL discovered that the sequence $L_n(f, B)$, belonging to the "most classical" matrix $B$ defined by

$$x_{\nu n} = 1 - \frac{2\nu}{n+1} \qquad (\nu = 1, 2, \ldots, n;\ n = 1, 2, \ldots),$$

can diverge in a whole subinterval of $[-1, +1]$ for such a simple function as $f(x) = \frac{1}{1+x^2}$. This would leave open the possibility, the situation can be saved by choosing another "better" $A$ matrix. But G. FABER[3] discovered in 1910 the shaking fact that *no* matrix $A$ can assure the uniform convergence of the polynomials $L_n(f, A)$ for every $f \in C$. His proof showed that it is essential for this phenomenon a property of the quantity

$$M_n(A) \equiv \max_{1 \leqq x \leqq +1} \lambda_n(x, A) = \max_{-1 \leqq x \leqq +1} \sum_{\nu=1}^{n} |l_{\nu n}(x, A)|, \tag{2.2}$$

namely that for *all* $A$-matrices we have

$$\overline{\lim_{n \to \infty}}\ M_n(A) = +\infty. \tag{2.3}$$

[2] $C$ denotes, as usual, the class of functions continuous for $-1 \leqq x \leqq +1$.

[3] G. FABER, Über die interpolatorische Darstellung stetiger Funktionen, *Jahresber. der Deutsch. Math. Ver.*, **23** (1914), pp. 190—210.

We call the quantity $\lambda_n(x, A)$ the $n^{th}$ Lebesgue function, the quantity $M_n(A)$ the $n^{th}$ Lebesgue constant belonging to the matrix $A$. Moreover, S. BERNSTEIN[4] proved the still more surprising fact that to every matrix $A$ there belongs an $x_0$ in $-1 \leqq x_0 \leqq +1$ and an $f_0 \in C$ with

(2.4) $$\overline{\lim_{n\to\infty}} |L_n(f_0, A)|_{x=x_0} = +\infty$$

what roots in the relation

(2.5) $$\overline{\lim_{n\to\infty}} \lambda_n(x_0, A) = +\infty.$$

**3.** These facts show that for the divergence theory of the Lagrange interpolation the functions $\lambda_n(x, A)$ are essential. But it was observed by FEJÉR[5] that these Lebesgue functions $\lambda_n(x, A)$ play also a role in the convergence theory. His simple reasoning reproduced for the sake of completeness in **9** gives the following theorem:

If the quantities $M_n(A)$ of (2.2) satisfy the inequality

(3.1) $$M_n(A) < c_1 n^\beta$$

with a fixed $0 < \beta < 1$ and numerical[6] $c_1$, then the polynomials $L_n(f, A)$ converge for $-1 \leqq x \leqq +1$ uniformly to $f(x)$ if[7] $f \in \mathrm{Lip}\,\gamma$, $\gamma > \beta$.

**4.** These results are responsible for the rather general opinion that the convergence-divergence theory of the Lagrange interpolation is by and large identical with the study of the Lebesgue constants $M_n(A)$. We have set ourselves the task to investigate to which extent this is true. We have found that going a little beyond the mere continuity this fails to be true; the result can quite vaguely expressed saying that *there is a "rough" and a "fine" convergence-divergence theory for the Lagrange interpolation*. To be more exact, let us consider, if

(4.1) $$0 < \beta < 1,$$

the class $A(\beta)$ of all $A$-matrices for which with arbitrarily small positive $\varepsilon$ we have

(4.2) $$\overline{\lim_{n\to\infty}} M_n(A) n^{-\beta-\varepsilon} < c_2(\varepsilon),$$

(4.3) $$\overline{\lim_{n\to\infty}} M_n(A) n^{-\beta+\varepsilon} > c_3(\varepsilon),$$

[4] S. BERNSTEIN, Sur la limitation des valeurs etc., *Bull. Acad. Sc. De l'URSS*, **8** (1931), pp. 1025—1050.

[5] L. FEJÉR, Lagrangesche Interpolation und die zugehörigen konjugierten Punkte, *Math. Ann.*, **106** (1932), pp. 1—55.

[6] Later on $c_2, c_3, \ldots$ denote generally also numerical constants. If some $c_\nu$ depends upon some parameters, the dependence will be explicitely stated.

[7] As usual, the class $\mathrm{Lip}\,\gamma$ denotes the totality of those functions which satisfy uniformly a Lipschitz-condition with the exponent $\gamma$ in $-1 \leqq x \leqq +1$.

i.e. for our matrices the Lebesgue constants $M_n(A)$ increase roughly speaking like $n^\beta$. We call for a fixed $\beta$ with (4. 1) the Lip $\gamma$ class of functions

| | |
|---|---|
| "a good class of functions $f(x)$ for $A(\beta)$", | if for all $A \in A(\beta)$ and $f \in$ Lip $\gamma$ $L_n(f, A)$ tend uniformly in $[-1, +1]$ to $f(x)$ for $n \to \infty$, |

respectively

| | |
|---|---|
| "a bad class of functions $f(x)$ for $A(\beta)$", | if for all $A \in A(\beta)$ there is an $f_1(x) \in$ Lip $\gamma$ such that $L_n(f_1, A)$ is unbounded in $[-1, +1]$ for $n \to \infty$. |

If a Lip $\gamma$ class is a good or a bad one, then its convergence-divergence behaviour is by and large determined by the numbers $M_n(A)$; one would be inclined to think that all Lip $\gamma$ classes with $\gamma < \beta$ are bad and all with $\gamma > \beta$ are good classes and thus the finer structure of the matrix $A$ cannot essentially influence the convergence-divergence behaviour for the respective Lip $\gamma$ class. A closer investigation showed, howewer, that this is not the case, there are values $\gamma$ depending only upon $\beta$ for which the Lip $\gamma$ classes are neither good nor bad ones. This means that the convergence-divergence behaviour is certainly not determined for the respective Lip $\gamma$ class alone by the Lebesgue constants $M_n(A)$; hence the convergence-divergence behaviour for the respective Lip $\gamma$ class depends upon the finer structure of $A$ and thus the determination of the convergence-divergence behaviour belongs to a "finer" theory. Thus even the existence of this "finer" theory is somewhat surprising; moreover we shall see that *for fixed* $0 < \beta < 1$ *all* Lip $\gamma$ *classes with*

(4. 4) $$\gamma < \frac{\beta}{\beta + 2}$$

*are bad classes, all* Lip $\gamma$ *classes with*

(4. 5) $$\gamma > \beta$$

*are good ones and the* Lip $\gamma$ *classes with*

(4. 6) $$\frac{\beta}{\beta + 2} < \gamma < \beta$$

*form the exact domain of the finer convergence-divergence theory.*

Similar questions arise in connection with orthogonal expansions, singular integrals, mechanical quadrature and generally with linear operations. Also other scales than the Lip-classes can be used. Further, the convergence behaviour can instead of uniform convergence refer e. g. to pointwise convergence. Finally, perhaps the matrix-class $A(\beta)$ can be defined more suitably than in (4. 2)—(4. 3).

**5.** Our result is given under (4. 4), (4. 5) and (4. 6). In order to prove it we brake the assertion into four parts. $\beta$ is fixed with

(5. 1) $$0 < \beta < 1.$$

a) If $\gamma < \frac{\beta}{\beta+2}$ and $A \in A(\beta)$, then there is an $f_2 \in \mathrm{Lip}\,\gamma$ such that the sequence $L_n(f_2, A)$ is unbounded for $-1 \leqq x \leqq 1$, i. e. the class Lip $\gamma$ is bad in this case.

b) If $\gamma > \beta$ and $A \in A(\beta)$, then the sequence $L_n(f, A)$ converges uniformly in $[-1, +1]$ to $f(x)$ whenever $f \in \mathrm{Lip}\,\gamma$, i. e. the class Lip $\gamma$ is good in this case.

c) If $\gamma > \frac{\beta}{\beta+2}$, then there is a matrix $A_0 \in A(\beta)$ such that the sequence $L_n(f, A_0)$ converges uniformly in $[-1, +1]$ to $f(x)$ whenever $f \in \mathrm{Lip}\,\gamma$, i. e. the class Lip $\gamma$ is certainly not a bad class.

d) If $\gamma < \beta$, then there is a matrix $A_1 \in A(\beta)$ and $f_3 \in \mathrm{Lip}\,\gamma$ such that the sequence $L_n(f_3, A_1)$ is unbounded for $-1 \leqq x \leqq +1$, i. e. the class Lip $\gamma$ is certainly not a good one.[8]

**6.** To prove the assertion a) we need the simple

LEMMA I. *If $A \in A(\beta)$, then we have for $\nu = 1, 2, \ldots, (n-1)$*

$$x_{\nu n} - x_{\nu+1, n} > c_4(\varepsilon) n^{-\beta-2-\varepsilon}. \tag{6. 1}$$

For the proof we consider the quantity

$$\frac{1}{x_{\nu n} - x_{\nu+1, n}}.$$

Owing to the definition of the fundamental functions and the mean-value theorem

$$\frac{1}{x_{\nu n} - x_{\nu+1, n}} = \frac{l_{\nu n}(x_{\nu n}, A) - l_{\nu n}(x_{\nu+1, n}, A)}{x_{\nu n} - x_{\nu+1, n}} = l'_{\nu n}(\zeta, A), \tag{6. 2}$$

where

$$x_{\nu+1, n} \leqq \zeta \leqq x_{\nu n}.$$

But using MARKOV's well-known theorem we have

$$|l'_{\nu n}(\zeta)| \leqq (n-1)^2 \max_{-1 \leqq x \leqq +1} |l_{\nu n}(x, A)|,$$

i. e. a fortiori

$$|l'_{\nu n}(\zeta, A)| < n^2 M_n(A). \tag{6. 3}$$

From (4. 2) we have $M_n(A) \leqq c_5(\varepsilon) n^{\beta+\varepsilon}$, i. e. (6. 2) and (6. 3) give

$$\frac{1}{x_{\nu n} - x_{\nu+1, n}} < c_5(\varepsilon) n^{\beta+2+\varepsilon}$$

which proves (6. 1)

[8] By modifications of the construction we could prove that in cases c) and d) also a matrix with $c_1 n^\beta \leqq \max_{-1 \leqq x \leqq +1} \sum_{\nu=1}^{n} |l_{\nu n}(x)| \leqq c_2 n^\beta$ could have been constructed with the other required properties. Also the investigation of the limiting cases is of interest.

**7.** Now we turn to the proof of the assertion a). Let

$$\sum_{\nu=1}^{n} |l_{\nu n}(x, A)|$$

maximal in $[-1, +1]$ for $x = \zeta_n$. With the convention

$$x_{0n} = -x_{n+1,n} = 1, \quad l_{0n}(x, A) \equiv l_{1n}(x, A), \quad l_{n+1,n}(x, A) \equiv l_{nn}(x, A)$$

we consider the function $\varphi_n(x)$ defined by the broken line with the vertices at the points

$$P_\nu \equiv (x_{\nu n}, \operatorname{sign} l_{\nu n}(\zeta_n, A)) \qquad (\nu = 0, 1, \ldots, n, n+1).$$

Then we have obviously

$$L_n(\varphi_n, A)_{x=\zeta_n} = \sum_{\nu=1}^{n} |l_{\nu n}(\zeta_n, A)| = M_n(A). \tag{7. 1}$$

According to the Lemma I all slopes of $\varphi_n(x)$ are absolutely

$$< c_5(\varepsilon) n^{\beta+2+\varepsilon}, \tag{7. 2}$$

and for $-1 \leqq x \leqq +1$

$$|\varphi_n(x)| \leqq 1. \tag{7. 3}$$

Now we can construct $f_2(x)$ in a way which is a suitably modified form of the resonance principle of LEBESGUE. Since $\gamma < \frac{\beta}{\beta+2}$, we may choose $\varepsilon$ so small that

$$0 < \varepsilon < \frac{\beta}{10} \tag{7. 4}$$

and

$$\frac{\beta - 2\varepsilon}{\beta + 2 + 2\varepsilon} > \gamma; \tag{7. 5}$$

we fix $\varepsilon$. According to (4. 3) there is an infinite sequence

$$2^2 \leqq n_1 < n_2 < \cdots$$

such that for $\nu = 1, 2, \ldots$

$$M_{n_\nu}(A) = \lambda_{n_\nu}(\zeta_{n_\nu}, A) > \frac{c_3(\varepsilon)}{2} n_\nu^{\beta-\varepsilon}. \tag{7. 6}$$

Now we select a suitable sub-sequence of the $n_\nu$'s which we shall denote by $r_\nu$'s. Let

$$r_1 = n_1$$

and we suppose

$$r_1, r_2, \ldots, r_{\nu-1}$$

are already defined. We distinguish two cases.

CASE I. Denoting

$$F_{\nu-1}(x) = \sum_{j=1}^{\nu-1} \frac{1}{r_j^{\beta-2\varepsilon}} \varphi_{r_j}(x)$$

the sequence

$$L_\mu(F_{\nu-1}, A) \qquad (\mu = 1, 2, \ldots)$$

is unbounded. In this case for

$$-1 \leqq x < x+h \leqq 1$$

we have

$$|F_{\nu-1}(x+h) - F_{\nu-1}(x)| \leqq \sum_{j=1}^{\nu-1} \frac{1}{r_j^{\beta-2\varepsilon}} |\varphi_{r_j}(x+h) - \varphi_{r_j}(x)|,$$

i. e. using (7. 2)

$$|F_{\nu-1}(x+h) - F_{\nu-1}(x)| \leqq c_5(\varepsilon) h \sum_{j=1}^{\nu-1} r_j^{2+3\varepsilon} < c_6(\varepsilon) h,$$

i. e. $F_{\nu-1}(x)$ belongs even to Lip 1.

CASE II. The sequence $L_\mu(F_{\nu-1}, A)$ is bounded, i. e. for $[-1, +1]$ and $\mu = 1, 2, \ldots$

(7. 7) $$|L_\mu(F_{\nu-1}, A)| \leqq C_{\nu-1}.$$

Then let $r_\nu$ be the smallest integer satisfying the conditions

(7. 8) $$r_\nu > r_{\nu-1}^2 \qquad (\text{i. e. also } > 2r_{\nu-1}),$$

(7. 9) $$r_\nu > C_{\nu-1}^{\frac{2}{\varepsilon}}.$$

We may suppose that we have for all $\nu$'s the Case II and we assert that in this case we may choose as $f_2(x)$ of assertion a)

(7. 10) $$f_2(x) = \sum_{j=1}^{\infty} \frac{1}{r_j^{\beta-2\varepsilon}} \varphi_{r_j}(x).$$

In order to show that $f_2(x) \in \mathrm{Lip}\,\gamma$ we write for an $x$ and $h$ satisfying

$$-1 \leqq x < x+h \leqq +1, \quad 0 < h < \frac{1}{100}$$

$$f_2(x+h) - f_2(x) = \sum_{j=1}^{p} + \sum_{j=p+1}^{\infty} \frac{1}{r_j^{\beta-2\varepsilon}} (\varphi_{r_j}(x+h) - \varphi_{r_j}(x))$$

where the index $p$ is defined uniquely by

(7. 11) $$r_p \leqq \left(\frac{1}{h}\right)^{\frac{1}{\beta+2+2\varepsilon}} < r_{p+1}.$$

Using for $\nu \leqq p$ (7. 2) and for $\nu \geqq p+1$ (7. 3), we get

$$|f_2(x+h) - f_2(x)| \leqq c_5(\varepsilon) h \sum_{j=1}^{p} r_j^{2+3\varepsilon} + 2 \sum_{j=p+1}^{\infty} r_j^{-\beta+2\varepsilon} <$$

$$< 2c_5(\varepsilon) \left\{ h p r_p^{2+3\varepsilon} + \frac{1}{r_{p+1}^{\beta-2\varepsilon}} \sum_{j=p+1}^{\infty} \left(\frac{r_{p+1}}{r_j}\right)^{\beta-2\varepsilon} \right\},$$

i. e. using (7. 8)

(7. 12) $$|f_2(x+h) - f_2(x)| \leqq c_6(\varepsilon, \beta) \left\{ h p r_p^{2+3\varepsilon} + \frac{1}{r_{p+1}^{\beta-2\varepsilon}} \right\}.$$

(7. 8) gives $r_p > 2^{p-1} r_1 \geqq 2^p$, i. e. $p < 2 \log r_p$; hence from this and (7. 11)

$$h p r_p^{2+3\varepsilon} < c_7(\varepsilon) h r_p^{2+4\varepsilon} < c_7(\varepsilon) h^{\frac{\beta-2\varepsilon}{\beta+2+2\varepsilon}}, \quad \frac{1}{r_{p+1}^{\beta-2\varepsilon}} < h^{\frac{\beta-2\varepsilon}{\beta+2+2\varepsilon}},$$

i. e. (7. 12) and (7. 5) give at once

$$|f_2(x+h) - f_2(x)| \leqq c_8(\varepsilon, \beta) h^{\frac{\beta-2\varepsilon}{\beta+2+2\varepsilon}} < c_8(\varepsilon, \beta) h^{\gamma}.$$

Thus our $f_2(x)$ belongs to the class Lip $\gamma$ as stated.

**8.** We have also to show that the polynomials $L_n(f_2, A)$ are unbounded in $[-1, +1]$. Let $s \geqq 2$ and write

$$f_2(x) = F_{s-1}(x) + \frac{1}{r_s^{\beta-2\varepsilon}} \varphi_{r_s}(x) + \Phi_s(x) \tag{8. 1}$$

with

$$\Phi_s(x) = \sum_{j=s+1}^{\infty} \frac{1}{r_j^{\beta-2\varepsilon}} \varphi_{r_j}(x).$$

(7. 7) gives for $-1 \leqq x \leqq +1$

$$|L_{r_s}(F_{s-1}, A)| \leqq C_{s-1}. \tag{8. 2}$$

From (7. 1) and (4. 3) we have

$$r_s^{-\beta+2\varepsilon} |L_{r_s}(\varphi_{r_s}, A)|_{x=\zeta_{r_s}} > c_9(\varepsilon) r_s^{\varepsilon}. \tag{8. 3}$$

Finally, (1. 8) gives for $-1 \leqq x \leqq +1$

$$|L_{r_s}(\Phi_s, A)| \leqq M_{r_s}(A) \max_{-1 \leqq x \leqq +1} |\Phi_s(x)|;$$

hence from (4. 2) and (7. 3), by (7. 8),

$$|L_{r_s}(\Phi_s, A)| \leqq c_{10}(\varepsilon) r_s^{\beta+\varepsilon} \sum_{j=s+1}^{\infty} \frac{1}{r_j^{\beta-2\varepsilon}} < c_{11}(\varepsilon, \beta) r_s^{-\beta+2\varepsilon}.$$

From this, (8. 2) and (8. 3) we obtained

$$|L_{r_s}(f_2, A)|_{x=\zeta_{r_s}} > c_9(\varepsilon) r_s^{\varepsilon} - C_{s-1} - c_{11}(\varepsilon, \beta) r_{s+1}^{-\beta+2\varepsilon}$$

what proves the unboundedness, using also (7. 9).

**9.** Next we turn to the proof of the assertion b) in **5.** This is based on an idea of S. BERNSTEIN adapted to interpolation by L. FEJÉR;[5] as told, we only reproduce it for the sake of completeness. Let $P_{n-1}(x)$ be the best-approximating polynomial of $(n-1)^{\text{th}}$ degree of $f(x)$ in $[-1, +1]$ in CHEBYSEV's sense. If $f \in \text{Lip}\, \gamma$, then according to S. BERNSTEIN we have in $[-1, +1]$

$$|f(x) - P_{n-1}(x)| \leqq c_{12} n^{-\gamma}. \tag{9. 1}$$

Owing to (1. 7) we have

$$L_n(P_{n-1}, A) \equiv P_{n-1}(x),$$

i. e.

$$L_n(f, A) - f(x) = L_n(f - P_{n-1}, A) + (P_{n-1}(x) - f(x))$$

and thus

$$|L_n(f, A)-f| \leqq c_{12} n^{-\gamma}+|L_n(f-P_{n-1}, A)|. \tag{9.2}$$

But using (1.3) and (9.1) we obtain

$$|L_n(f-P_{n-1}, A)| \leqq c_{12} n^{-\gamma} \sum_{\nu=1}^{n} |l_{\nu n}(x, A)|. \tag{9.3}$$

Choosing in (4.2) $\varepsilon$ so small that $\beta+\varepsilon<\gamma-\varepsilon$ and fixing it, we obtain

$$\sum_{\nu=1}^{n} |l_{\nu n}(x, A)| \leqq c_{13}(\varepsilon) n^{\beta+\varepsilon} < c_{13}(\varepsilon) n^{\gamma-\varepsilon}$$

and from (9.3)

$$|L_n(f-P_{n-1}, A)| \leqq c_{14}(\varepsilon) n^{-\varepsilon}.$$

This and (9.2) prove already the assertion b).

**10.** Next we turn to the proof of the assertion c) in **5.** We shall show that the matrix whose $n^{\text{th}}$ row consists of

$$x_{1n} = \cos\left(\frac{3\pi}{2n}-\frac{1}{n^{1+\beta}}\right), \quad x_{\nu n} = \cos\frac{2\nu-1}{2n}\pi \qquad (\nu=2, 3, \ldots, n) \tag{10.1}$$

belongs to $A(\beta)$ and fulfils the requirements for $A_0$ of the assertion c). In what follows, we shall speak about one line of the matrix so that instead of $x_{\nu n}$ and $l_{\nu n}(x, A_0)$ we may write $x_\nu$ and $l_\nu(x, A_0)$. We have obviously

$$x_1-x_2 = 2\sin\frac{1}{2n^{1+\beta}}\sin\left(\frac{3\pi}{2n}-\frac{1}{2n^{1+\beta}}\right) = \frac{3\pi}{2} n^{-\beta-2}+O(n^{-2-2\beta}), \tag{10.2}$$

i. e. for $n>n_0$

$$x_1-x_2 < 2\pi n^{-\beta-2}. \tag{10.3}$$

Hence

$$\frac{n^{\beta+2}}{2\pi} < \frac{1}{x_1-x_2} = \frac{l_1(x_1, A_0)-l_1(x_2, A_0)}{x_1-x_2} = \left|\frac{l_1(x_1, A_0)-l_1(x_2, A_0)}{x_1-x_2}\right| \leqq$$

$$\leqq \max_{-1\leqq x\leqq +1}\left|\frac{d}{dx} l_1(x, A_0)\right| < n^2 \max_{-1\leqq x\leqq 1} |l_1(x, A_0)| \leqq n^2 M_n(A_0),$$

using again MARKOV's theorem. Hence

$$M_n(A_0) \geqq \frac{1}{2\pi} n^{\beta}. \tag{10.4}$$

If we show that

$$M_n(A_0) \leqq c_{15} n^{\beta}, \tag{10.5}$$

then $A_0$ belongs indeed to $A(\beta)$.

**11.** For the proof of (10.5) we need some lemmas. We shall denote by $T$ the matrix the $n^{th}$ row of which consists of the numbers

$$x_\nu^* = \cos\frac{2\nu-1}{2n}\pi \qquad (\nu=1,2,\ldots,n) \tag{11.1}$$

and by $l_\nu(x,T)$ the fundamental functions belonging to $T$. Then we need the

LEMMA II. *We have for* $-1 \leqq x \leqq +1$

$$\sum_{\nu=1}^{n} |l_\nu(x,T)| \leqq c_{16}\log n.$$

This lemma is well known.[9] We need further the

LEMMA III. *If* $n \geqq 4$, *we have for* $-1 \leqq x \leqq +1$

$$\sum_{\nu=3}^{n} |l_\nu(x,A_0)| \leqq c_{17}\log n.$$

PROOF. For $\nu \geqq 2$ we have

$$\frac{l_\nu(x,A_0)}{l_\nu(x,T)} = \frac{x-x_1}{x_\nu-x_1}\cdot\frac{x_\nu-x_1^*}{x-x_1^*} = \left(1-\frac{x_1^*-x_1}{x_1^*-x}\right)\left(1+\frac{x_1^*-x_1}{x_1-x_\nu}\right). \tag{11.2}$$

Since for $\nu \geqq 3$, $n \geqq 3$

$$0 < 1+\frac{x_1^*-x_1}{x_1-x_\nu} < 1+\frac{\cos\frac{\pi}{2n}-\cos\frac{3\pi}{2n}}{\cos\frac{3\pi}{2n}-\cos\frac{5\pi}{2n}} = 1+\frac{1}{2\cos\frac{\pi}{n}} \leqq 1+\frac{1}{\sqrt{2}},$$

we have

$$\left|\frac{l_\nu(x,A_0)}{l_\nu(x,T)}\right| \leqq \left(1+\frac{1}{\sqrt{2}}\right)\left|1-\frac{x_1^*-x_1}{x_1^*-x}\right|. \tag{11.3}$$

If $-1 \leqq x \leqq x_1$, then owing to

$$0 < 1-\frac{x_1^*-x_1}{x_1^*-x} < 1$$

we have from (11.3) for $-1 \leqq x \leqq x_1$

$$|l_\nu(x,A_0)| \leqq 2|l_\nu(x,T)|,$$

i. e. summing for $\nu=3,4,\ldots,n$ and using Lemma II we obtain

$$\sum_{\nu=3}^{n} |l_\nu(x,A_0)| \leqq 2c_{17}\log n, \tag{11.4}$$

i. e. Lemma III is proved for $-1 \leqq x \leqq x_1$. To prove it also for $x_1 \leqq x \leqq 1$ we write (11.3) in the form

$$|l_\nu(x,A_0)| < 2\left|\frac{x-x_1}{x-x_1^*}\right||l_\nu(x,T)| = 2|x-x_1|\left|\frac{l_\nu(x,T)}{x-x_1^*}\right|.$$

[9] See S. N. BERNSTEIN, Quelques remarques sur l'interpolation, *Math. Ann.*, **79** (1918), pp. 1—12.

Let us observe that for $\nu \geqq 3$ and $x_1 \leqq x \leqq x_1^*$ we have

(11.5) $$\operatorname{sign} l_\nu(x, T) = (-1)^\nu$$

and for $x_1^* \leqq x \leqq 1$

(11.6) $$\operatorname{sign} l_\nu(x, T) = (-1)^{\nu+1}.$$

Hence for $x_1 \leqq x \leqq x_1^*$, $\nu \geqq 3$

$$|l_\nu(x, A_0)| < 2|x-x_1| \frac{(-1)^\nu l_\nu(x, T)}{|x-x_1^*|},$$

i. e.

(11.7) $$\sum_{\nu=3}^{n} |l_\nu(x, A_0)| < 2|x-x_1| \left| \frac{\sum_{\nu=3}^{n} (-1)^\nu l_\nu(x, T)}{x-x_1^*} \right|.$$

A similar reasoning shows, (11.7) holds also for $x_1^* \leqq x \leqq 1$, i. e. (11.7) holds for $x_1 \leqq x \leqq 1$. Applying the mean-value theorem to the expression

$$\frac{\sum_{\nu=3}^{n} (-1)^\nu l_\nu(x, T)}{x-x_1^*}$$

and also

$$|x-x_1| < 1-\cos\frac{3\pi}{2n} < \frac{10^2}{n^2},$$

we obtain from (11.7) for $x_1 \leqq x \leqq 1$

$$\sum_{\nu=3}^{n} |l_\nu(x, A_0)| < 2\frac{10^2}{n^2} \max_{-1\leqq x\leqq +1} \left| \frac{d}{dx} \sum_{\nu=3}^{n} (-1)^\nu l_\nu(x, T) \right|.$$

Using again MARKOV's theorem the right side is

$$< \frac{200}{n^2}(n-1)^2 \max_{-1\leqq x\leqq +1} \left| \sum_{\nu=3}^{n} (-1)^\nu l_\nu(x, T) \right| < 200 \max_{-1\leqq x\leqq 1} \sum_{\nu=1}^{n} |l_\nu(x, T)| \leqq$$

$$\leqq 200\, c_{16} \log n$$

by Lemma II. By this and (11.4) Lemma III is proved.

**12.** It follows from Lemma III that (10.5) will be proved if we succeed in showing that for $-1 \leqq x \leqq +1$ we have

(12.1) $$|l_1(x, A_0)| \leqq c_{18}\, n^\beta, \quad |l_2(x, A_0)| \leqq c_{18}\, n^\beta.$$

This will follow as a byproduct from the following Lemma which we shall need in **13.**

LEMMA IV. *For $-1 \leqq x \leqq +1$ we have*

$$|l_1(x, A_0) + l_2(x, A_0)| \leqq c_{19}.$$

PROOF. From (1.4) we have

(12.2) $$\frac{l_1(x, A_0)}{l_1(x, T)} = \prod_{j=2}^{n} \frac{x_1^*-x_j}{x_1-x_j} = \frac{T_n'(x_1^*)}{T_n'(x_2)} \left\{ (x_1-x_1^*) \frac{T_n'(x_2)}{T_n(x_1)} \right\}.$$

The factor in the bracket is

$$(12.3)\quad \left\{\cos\frac{\pi}{2n}-\cos\left(\frac{3\pi}{2n}-\frac{1}{n^{1+\beta}}\right)\right\}\left\{-\frac{n}{\sin\frac{3\pi}{2n}}\cdot\frac{1}{\cos\left(\frac{3\pi}{2}-n^{-\beta}\right)}-2\,\frac{\pi}{2n}\left\{1+O\left(\frac{1}{n^{\beta}}\right)\right\}\right\}\cdot$$
$$\cdot\frac{\pi}{n}\left\{1+O\left(\frac{1}{n^{\beta}}\right)\right\}\left\{\frac{2}{3\pi}n^{2}\left\{1+O\left(\frac{1}{n^{2}}\right)\right\}\left\{n^{\beta}\left\{1+O\left(\frac{1}{n^{2\beta}}\right)\right\}\right\}=-\frac{2\pi}{3}n^{\beta}+O(1).$$ [10]

Taking in account that

$$\frac{x_2-x_1^*}{x_2-x_1}=\frac{\cos\frac{3\pi}{2n}-\cos\frac{\pi}{2n}}{\cos\frac{3\pi}{2n}-\cos\left(\frac{3\pi}{2n}-n^{-1-\beta}\right)}=\frac{\pi^2 n^{-2}\left\{1+O\left(\frac{1}{n^2}\right)\right\}}{\frac{3\pi}{2}n^{-2-\beta}\{1+O(n^{-\beta})\}}=$$
$$=\frac{2\pi}{3}n^{\beta}\{1+O(n^{-\beta})\}=\frac{2\pi}{3}n^{\beta}+O(1),$$

this and (12. 3) means that

$$(x_1-x_1^*)\frac{T_n'(x_2)}{T_n(x_1)}=-\frac{x_2-x_1^*}{x_2-x_1}+O(1)$$

and from (12. 2), using the explicit form of $l_1(x, T)$,

$$(12.4)\quad \begin{aligned} l_1(x, A_0)&=\frac{T_n(x)}{x-x_1^*}\cdot\frac{1}{T_n'(x_2)}\left\{-\frac{x_2-x_1^*}{x_2-x_1}+O(1)\right\}=\\ &=-\frac{x_2-x_1^*}{x_2-x_1}\cdot\frac{T_n(x)}{(x-x_1^*)T_n'(x_2)}+O(1)|l_1(x, T)|\left|\frac{T_n'(x_1^*)}{T_n'(x_2)}\right|. \end{aligned}$$

As well known,[5] we have for $-1\leqq x\leqq +1$

$$(12.5)\qquad |l_\nu(x, T)|\leqq\sqrt{2}\qquad (\nu=1, \ldots, n),$$

i. e. the second term on the right of (12. 4) is

$$O(1)\left|\frac{T_n'(x_1^*)}{T_n'(x_2)}\right|=O(1)\frac{\sin\frac{3\pi}{2n}}{\sin\frac{\pi}{2n}}=O(1);$$

thus from (12. 4)

$$(12.6)\qquad l_1(x, A_0)=-\frac{x_2-x_1^*}{x_2-x_1}\cdot\frac{T_n(x)}{x-x_1^*}\cdot\frac{1}{T_n'(x_2)}+O(1).$$

Further we have

$$\frac{l_2(x, A_0)}{l_2(x, T)}=\frac{x-x_1}{x-x_1^*}\cdot\frac{x_2-x_1^*}{x_2-x_1},$$

[10] The $O$-sign refers to $n\to\infty$ but always uniformly for $-1\leqq x\leqq +1$.

i. e. using the explicit form of $l_2(x, T)$

$$l_2(x, A_0) = \frac{x_2 - x_1^*}{x_2 - x_1} \cdot \frac{x - x_1}{x - x_1^*} \cdot \frac{T_n(x)}{x - x_2} \cdot \frac{1}{T_n'(x_2)}.$$

From this and (12. 6)

$$|l_1(x, A_0) + l_2(x, A_0)| = O(1) + \left|\frac{x_2 - x_1^*}{x_2 - x_1}\right| \left|\frac{T_n(x)}{T_n'(x_2)}\right| \left| -\frac{1}{x - x_1^*} + \frac{x - x_1}{x - x_1^*} \cdot \frac{1}{x - x_2}\right| =$$

(12. 7) $$= O(1) + \left|\frac{x_2 - x_1^*}{x_2 - x_1}\right| \left|\frac{T_n(x)}{T_n'(x_2)\,(x - x_1^*)}\right| \left|\frac{x_1 - x_2}{x - x_2}\right| =$$

$$= O(1) + |x_2 - x_1^*| \frac{1}{|T_n'(x_2)|} \left|\frac{T_n(x)}{(x - x_1^*)(x - x_2)}\right| = O(1) + O\left(\frac{1}{n^4}\right) \left|\frac{T_n(x)}{(x - x_1^*)(x - x_2)}\right|.$$

Since we have for $-1 \leqq x \leqq \cos \frac{2\pi}{n}$

$$\left|\frac{T_n(x)}{(x - x_1^*)(x - x_2)}\right| \leqq \frac{1}{\left(\cos \frac{\pi}{2n} - \cos \frac{2\pi}{n}\right)\left(\cos \frac{3\pi}{2n} - \cos \frac{2\pi}{n}\right)} = O(n^4)$$

for $\cos \frac{2\pi}{n} \leqq x \leqq \cos \frac{\pi}{n}$,

$$\left|\frac{T_n(x)}{(x - x_1^*)(x - x_2)}\right| \leqq \frac{1}{\cos \frac{\pi}{2n} - \cos \frac{\pi}{n}} \cdot \max_{-1 \leqq x \leqq +1} \left|\frac{T_n(x)}{x - x_2}\right| = O(n^4)$$

and similarly for $\cos \frac{\pi}{n} \leqq x \leqq +1$, Lemma IV follows from (12. 7) indeed.

In this proof is at the same time a proof for (12. 1) contained. We may write namely (12. 6) in the form

$$l_1(x, A_0) = -\frac{x_2 - x_1^*}{x_2 - x_1} \frac{T_n'(x_1^*)}{T_n'(x_2)} l_1(x, T) + O(1) = \frac{x_2 - x_1^*}{x_2 - x_1} \cdot \frac{\sin \frac{3\pi}{2n}}{\sin \frac{\pi}{2n}} l_1(x, T) + O(1).$$

Taking in account (12. 5), further

$$\frac{\sin \frac{3\pi}{2n}}{\sin \frac{\pi}{2n}} = O(1)$$

and

$$\frac{x_2 - x_1^*}{x_2 - x_1} = O(n^\beta),$$

the relation $l_1(x, A_0) = O(n^\beta)$ follows indeed. For $l_2(x, A_0)$ the assertion follows using Lemma IV. Thus our matrix belongs to $A(\beta)$ indeed.

**13.** In order to finish the proof of the assertion c) in **5** we have to show that the polynomials $L_n(f, A_0)$ converge uniformly in $[-1, +1]$ to $f(x)$ whenever $f \in \text{Lip}\,\gamma$, $\gamma > \frac{\beta}{\beta+2}$. Let $\eta$ be so small that

(13. 1) $$\gamma > \frac{\beta}{\beta+2} + \eta$$

and fixed. Let further

(13. 2) $$\left[\frac{n}{\log n}\right] = m$$

and $P_m(x)$ should stand for the best-approximating polynomial of $f(x)$ in $[-1, +1]$ in CHEBYSEV's sense. Then we have again

$$L_n(P_m, A_0) \equiv P_m(x)$$

and proceeding exactly as in **9** we obtain, using the abbreviation

(13. 3) $$f(x_\nu) - P_m(x_\nu) = y_\nu,$$

that

(13. 4) $$L_n(f, A_0) - f = (P_m - f) + \sum_1^n y_\nu l_\nu(x, A_0).$$

Using again S. BERNSTEIN's theorem mentioned in **9** we have

(13. 5) $$|f - P_m| \leqq c_{20} m^{-\gamma} < c_{21} n^{-\gamma} \log^\gamma n,$$

i. e. from (13. 4) we get

(13. 6) $$|L_n(f, A_0) - f| \leqq c_{21} n^{-\gamma} \log^\gamma n + |y_1 l_1(x, A_0) + y_2 l_2(x, A_0)| + \sum_{\nu=3}^n |y_\nu| |l_\nu(x, A_0)|.$$

Taking in account (13. 5) and Lemma III the last sum is $O(m^{-\gamma} \log n) = O(n^{-\gamma} \log^{\gamma+1} n)$; hence from (13.6) we have

(13. 7) $$|L_n(f, A_0) - f| \leqq$$
$$\leqq c_{21} n^{-\gamma} \log^{\gamma+1} n + |y_1| \, |l_1(x, A_0) + l_2(x, A_0)| + |y_1 - y_2| |l_2(x, A_0)|.$$

Taking in account Lemma IV and (12. 1), (13. 7) gives

(13. 8) $$|L_n(f, A_0) - f| \leqq c_{22} \{ n^{-\gamma} \log^{\gamma+1} n + n^\beta |y_1 - y_2| \}.$$

For the estimation of $|y_1 - y_2|$ we write

(13. 9) $$|y_1 - y_2| \leqq |f(x_1) - f(x_2)| + |P_m(x_1) - P_m(x_2)|.$$

Since $f \in \text{Lip}\,\gamma$ and

(13. 10) $$|x_1 - x_2| = \cos\left(\frac{3\pi}{2n} - \frac{1}{n^{1+\beta}}\right) - \cos\frac{3\pi}{2n} < \frac{100}{n^{2+\beta}},$$

we have, using (13. 1),

(13. 11) $$|f(x_1) - f(x_2)| \leqq c_{23} \left(\frac{100}{n^{2+\beta}}\right)^\gamma \leqq c_{24} n^{-(2+\beta)\left(\frac{\beta}{\beta+2}+\eta\right)} = c_{24} n^{-\beta} n^{-(\beta+2)\eta}.$$

Further, using again MARKOV's theorem and (13. 10),

$$|P_m(x_1)-P_m(x_2)| \leqq \frac{100}{n^{2+\beta}} \max_{-1\leqq x\leqq +1} |P_m'(x)| \leqq \frac{100 m^2}{n^{2+\beta}} \max_{-1\leqq x\leqq +1} |P_m(x)| <$$
$$< \frac{100}{n^\beta \log^2 n}(1+\max_{-1\leqq x\leqq +1} |f(x)|).$$

From this, (13.8) and (13. 11) it follows that the polynomials $L_n(f, A_0)$ converge in $[-1, +1]$ uniformly to $f(x)$, i. e. the proof of the assertion c) is finished.

**14.** Finally we shall prove the assertion d) in **5.** Since the proof follows the line of that of a) we shall not go into all details. It is obviously sufficient to define $A_1$ only for even $n=2k$. With our $\beta$ we define

$$\vartheta=\frac{\beta^2}{2}\left(\frac{\log k}{k}\right)^2 \tag{14. 1}$$

and the $2k^{th}$ line of our matrix $A_1$ should be given by the zeros $\zeta_1,\ldots,\zeta_{2k}$ of

$$\omega_{2k}(x, A_1)\equiv \omega_{2k}(x)\equiv T_k(1+\vartheta-(2+\vartheta)x^2)=0 \quad (T_k(\cos\vartheta)=\cos k\vartheta), \tag{14. 2}$$

i. e. with $x_\nu=\cos\dfrac{2\nu-1}{2k}\pi$

$$\begin{aligned}\zeta_\nu &= \sqrt{\frac{1+\vartheta-x_\nu}{2+\vartheta}} \qquad (\nu=1,2,\ldots,k),\\ \zeta_{-\nu} &= -\sqrt{\frac{1+\vartheta-x_\nu}{2+\vartheta}} \qquad (\nu=1,2,\ldots,k).\end{aligned} \tag{14. 3}$$

We have obviously

$$\left(\frac{\sqrt{\vartheta}}{2}<\right)\sqrt{\frac{\vartheta}{2+\vartheta}}<\zeta_1<\zeta_2<\cdots<\zeta_k<1. \tag{14. 4}$$

Since for $1\leqq|\nu|\leqq k$

$$|\omega_{2k}'(\zeta_\nu)|=\left|\frac{dT_k(u)}{du}\right|_{u=1+\vartheta-(2+\vartheta)\zeta_\nu^2}\cdot 2(2+\vartheta)|\zeta_\nu|=\frac{k}{\sqrt{1-x_{|\nu|}^2}}2(2+\vartheta)|\zeta_\nu|,$$

we get — denoting the fundamental functions by $l_{\pm\nu}(x, A_1)$ $(\nu=1,2,\ldots,k)$ —

$$\lambda_{2k}(x, A_1)=\sum_{1\leqq|\nu|\leqq k}|l_\nu(x, A_1)|=\frac{1}{2(2+\vartheta)k}\sum_{1\leqq|\nu|\leqq k}\frac{\sqrt{1-x_{|\nu|}^2}}{|\zeta_\nu|}\frac{|\omega_{2k}(x)|}{|x-\zeta_\nu|}. \tag{14. 5}$$

**15.** First we are going to show (by rough estimations)

$$\sum_{|\nu|=1}^{k}|l_\nu(x, A_1)|<c_{25}(\beta)\,k^\beta\log^2 k. \tag{15. 1}$$

We may obviously suppose $0\leqq x\leqq 1$. Since for all of our $\nu$'s we have

$$\frac{\sqrt{1-x_\nu}}{|\zeta_\nu|}=\sqrt{2+\vartheta}\sqrt{\frac{1-x_\nu}{1+\vartheta-x_\nu}}<2$$

and

$$\sqrt{1+x_\nu} < 10\,\frac{k-\nu+1}{k},$$

we have

(15. 2) $$M_{2k}(A_1) < \frac{4}{k^2} \sum_{1 \le |\nu| \le k} \frac{k-|\nu|+1}{|x-\zeta_\nu|}\,|\omega_{2k}(x)|.$$

What can be said on $(\zeta_{\nu+1}-\zeta_\nu)$? For

$$k \geqq |\nu| \geqq \frac{k}{2}$$

we have from (14. 3)

(15. 3) $$|\zeta_{\nu+1}-\zeta_\nu| = \frac{1}{\sqrt{2+\vartheta}} \cdot \frac{x_\nu - x_{\nu+1}}{\sqrt{1+\vartheta-x_\nu}+\sqrt{1+\vartheta-x_{\nu+1}}} > \frac{k-|\nu|+1}{20k^2}.$$

For

$$\frac{k}{2} > |\nu| \geqq 1$$

we have from (15. 3)

(15. 4) $$|\zeta_{\nu+1}-\zeta_\nu| > c_{26}(\beta)\frac{\dfrac{\nu}{k^2}}{\sqrt{\dfrac{\nu^2}{k^2}+\dfrac{\log^2 k}{k^2}}} > \frac{c_{27}(\beta)}{k\log k}.$$

Let first be

$$0 \leqq x \leqq \zeta_1.$$

Then we have from (15. 2)

(15. 5) $$\lambda_{2k}(x, A_1) < \frac{8}{k^2}|\omega_{2k}(x)| \sum_{\nu=1}^{k} \frac{k-\nu+1}{\zeta_\nu - x}.$$

Owing to (15. 4) we have

$$\sum_{\nu=1}^{k} \frac{k-\nu+1}{\zeta_\nu - x} \leqq \frac{k}{\zeta_1 - x} + k \sum_{2 \le \nu < \frac{k}{2}} \frac{1}{\zeta_\nu - x} + k \sum_{\frac{k}{2} \le \nu \le k} \frac{1}{\zeta_\nu - x} <$$

$$< \frac{k}{\zeta_1 - x} + c_{28}k^2 + k \sum_{2 \le \nu < \frac{k}{2}} \frac{1}{(\nu-1)c_{27}(\beta)\dfrac{1}{k\log k}} < \frac{k}{\zeta_1 - x} + c_{29}(\beta)k^2\log^2 k,$$

i. e. from (15. 5) for such values $x$

(15. 6) $$\lambda_{2k}(x, A_1) < c_{30}(\beta)\left\{\frac{1}{k}\left|\frac{\omega_{2k}(x)}{x-\zeta_1}\right| + |\omega_{2k}(x)|\log^2 k\right\}.$$

But for $0 \leqq x \leqq \xi_1$ owing to (14. 1) we have for $k > c_{31}$

(15. 7) $$\begin{aligned}|\omega_{2k}(x)| &\leqq T_k(1+\vartheta) < \frac{1}{2}\{1+(1+\vartheta+\sqrt{(1+\vartheta)^2-1})^k\} < \\ &< (1+\vartheta+\sqrt{2\vartheta+\vartheta^2})^k < e^{k\left(\vartheta+\sqrt{2\vartheta}\sqrt{1+\frac{\vartheta}{2}}\right)} < 2e^{k\sqrt{2\vartheta}} = 2k^\beta.\end{aligned}$$

Further, taking in account that for $0 \leqq x \leqq \zeta_1$

$$\left|\frac{d\omega_{2k}(x)}{dx}\right| = 2(2+\vartheta)|x| \left|\frac{dT_k(u)}{du}\right|_{u=1+\vartheta-(2+\vartheta)x^2} =$$

$$= k(2+\vartheta)|x| \left\{ (u+\sqrt{u^2-1})^{k-1}\left(1+\frac{u}{\sqrt{u^2-1}}\right) + \right.$$

$$\left. + (u-\sqrt{u^2-1})^{k-1}\left(1-\frac{u}{\sqrt{u^2-1}}\right)\right\}_{u=1+\vartheta-(2+\vartheta)x^2} =$$

$$= k(2+\vartheta)|x| \left\{ \frac{(u+\sqrt{u^2-1})^k-(u-\sqrt{u^2-1})^k}{\sqrt{u^2-1}} \right\}_{u=1+\vartheta-(2+\vartheta)x^2} \leqq$$

$$\leqq k(2+\vartheta)\zeta_1 \left\{ \quad \right\}_{u \quad 1+\vartheta} < \frac{k(2+\vartheta)}{\sqrt{2\vartheta}} \sqrt{\frac{2\sin^2\frac{\pi}{4k}+\vartheta}{2}} \cdot 2k^\beta < 10k^{\beta+1},$$

we have for $0 \leqq x \leqq \zeta_1$

$$(15.8) \qquad \frac{1}{k}\left|\frac{\omega_{2k}(x)}{x-\zeta_1}\right| \leqq \frac{1}{k} \max_{0 \leqq x \leqq \zeta_1} |\omega_{2k}'(x)| < 10k^\beta.$$

(15.6), (15.7) and (15.8) give at once (15.1) for $0 \leqq x \leqq \xi_1$. For $\zeta_1 \leqq x \leqq 1$ the proof of (15.1) runs on similar lines and also a much better estimation of $\lambda_{2k}(x, A_1)$ could have been given; we omit the details. Thus (15.1) is proved.

**16.** Next we show that

$$(16.1) \qquad \lambda_{2k}(0, A_1) > c_{32}(\beta)\, k^\beta.$$

Since for $k > c_{33}(\beta)$

$$|\omega_{2k}(0)| = T_k(1+\vartheta) > \frac{1}{2}\,(1+\vartheta+\sqrt{2\vartheta+\vartheta^2})^k > \frac{1}{4}\,k^\beta,$$

we obtain from (14.5)

$$\lambda_{2k}(0, A_1) > \frac{k^{\beta-1}}{4(2+\vartheta)} \sum_{\nu=1}^{k} \frac{\sqrt{1-x_\nu^2}}{\zeta_\nu^2} > \frac{k^{\beta-1}}{4} \sum_{\frac{k}{10} \leqq \nu \leqq \frac{9}{10}k} \frac{\sqrt{1-x_\nu^2}}{1+\vartheta-x_\nu} > c_{33}(\beta)\, k^\beta.$$

This and (15.1) show that our matrix $A_1$ belongs indeed to $A(\beta)$.

**17.** What we actually proved, is a little more and this is what we shall need. We showed

$$\sum_{1 \leqq |\nu| \leqq \frac{1}{2}k} |l_\nu(0, A_1)| > c_{34}(\beta)\, k^\beta,$$

i. e.

$$(17.1) \qquad \sum_{1 \leqq |\nu| \leqq \frac{1}{2}k} |l_\nu(\Theta_{2k}, A_1)| \equiv \max_{-\frac{1}{4} \leqq x \leqq \frac{1}{4}} \sum_{1 \leqq |\nu| \leqq \frac{1}{2}k} |l_\nu(x, A_1)| \geqq c_{35}(\beta)\, k^\beta.$$

Having this, the construction of $f_3(x)$ runs as follows. Introducing the notation

(17. 2)
$$y_{\nu,2k}=\begin{cases} \operatorname{sign} l_{\nu,2k}(\Theta_{2k}, A_1) & \text{for } 1\leqq|\nu|\leqq\frac{k}{2}, \\ 0 & \text{for } \frac{k}{2}<|\nu|\leqq k, \end{cases}$$

we consider the function $\psi_{2k}(x)$ defined by the broken line with the vertices at the points

$$P_\nu^*\equiv(\zeta_{\nu,2k},\, y_{\nu,2k}) \qquad (1\leqq|\nu|\leqq k);$$

we use also here the convention of **7** to complete the graph of $\psi_{2k}(x)$ in the whole interval $[-1, +1]$ by affixing two horizontals at the ends. We have obviously

(17. 3)
$$L_{2k}(\psi_{2k}, A_1)_{x=\Theta_{2k}}=\sum_{1\leqq|\nu|\leqq\frac{k}{2}}|l_{\nu,2k}\,\Theta_{2k}, A_1)|>c_{35}(\beta)k^{\beta}.$$

According to (15. 4) the slopes of $\psi_{2k}(x)$ are absolutely

$$\leqq\frac{1}{c_{27}(\beta)}k\log k,$$

for $-\frac{51}{100}\leqq x\leqq\frac{51}{100}$ and $k\geqq c_{36}(\beta)$ we have

(17. 4)
$$|\psi_{2k}(x)|\leqq 1,$$

for $\frac{51}{100}\leqq|x|\leqq 1$ ($x$ real)

$$\psi_{2k}(x)=0$$

and for $-1\leqq x<x+h\leqq+1$

(17. 5)
$$|\psi_{2k}(x+h)-\psi_{2k}(x)|\leqq c_{37}(\beta)hk\log k.$$

Now our function $f_3(x)$ will be of the form

$$f_3(x)=\sum_{\nu=1}^{\infty}\frac{1}{r_\nu^{\beta-\varepsilon}}\psi_{2r_\nu}(x)$$

with sufficiently quickly increasing $r_\nu$-indices, as in **7.** The proof of unboundedness of $L_n(f_3, A_1)$ needs only slight changes compared to that of **8,** so we can omit the details. As to the Lipschitz exponent of $f_3(x)$ we have ($p$ to be disposed later)

$$|f_3(x+h)-f_3(x)|\leqq\sum_{\nu\leqq p}\frac{1}{r_\nu^{\beta-\varepsilon}}|\psi_{2r_\nu}(x+h)-\psi_{2r_\nu}(x)|+$$

$$+\sum_{\nu\geqq p+1}\frac{1}{r_\nu^{\beta-\varepsilon}}(|\psi_{2r_\nu}(x+h)|+|\psi_{2r_\nu}(x)|)$$

or owing to (17.5) and (17.4) "essentially" (owing to the quick increase of the $r_\nu$'s)

$$|f_3(x+h)-f_3(x)| \leqq c_{38}(\beta) \left\{ h r_p^{1-\beta+2\varepsilon} + \frac{1}{r_{p+1}^{\beta-\varepsilon}} \right\}.$$

Choosing $p$ so that

$$r_p \leqq \frac{1}{h} < r_{p+1},$$

everything follows since $\varepsilon$ is arbitrary small. Hence assertion d) is also proved.

(*Received 22 January 1955*)

## РОЛЬ ФУНКЦИЙ ЛЕБЕГА В ТЕОРИИ ИНТЕРПОЛЯЦИИ ЛАГРАНЖА

П. Ердёш (Иерусалим) и П. Туран (Будапешт)

(Резюме)

Главным результатом настоящей работы является доказательство того, что в теории сходимости и расходимости интерполяции Лагранжа следует различать между „грубой" и „тонкой" теориями. „Грубой" является часть теории, зависящая только от быстроты возрастания „лебеговых констант"

$$M_n(A) = \max_{-1 \leqq x \leqq +1} \sum_{\nu=1}^{n} |l_\nu(x)|,$$

где $A$ означает основную матрицу интерполяции, „тонкие" же другие результаты. Главная задача состояла в отделении этих двух теорий. Из различных возможных точек зрения мы здесь остановимся на той, при которой матрицы $A$ для произвольно малого положительного числа $\varepsilon$, и для фиксированного $0 < \beta < 1$ подчинены условиям

$$\text{(1)} \qquad \overline{\lim_{n\to\infty}} \frac{M_n(A)}{n^{\beta+\varepsilon}} < c_1(\varepsilon) \, (< \infty),$$

$$\text{(2)} \qquad \underline{\lim_{n\to\infty}} \frac{M_n(A)}{n^{\beta-\varepsilon}} > c_2(\varepsilon) \, (> 0),$$

а следует определять те классы функций $\mathrm{Lip}\,\alpha$ в $[-1,+1]$, для которых при любой матрице удовлетворяющей условиям (1) и (2) интерполяция пригодна, и те классы, для которых интерполяция при любых таких матрицах непригодна. Оказывалось, что если исходить из классы матриц удовлетворяющих (1) и (2), то „тонкая" теория есть теория сходимости-расходимости тех класс $\mathrm{Lip}\,\alpha$, для которых

$$\frac{\beta}{\beta+2} < \alpha < \beta.$$

Более точно, имеет место следующее:

a) Если для $A$ справедливы (1) и (2) и имеет место $\alpha < \frac{\beta}{\beta+2}$, то существует функция $f_1 \in \mathrm{Lip}\,\alpha$, для которой взятые по $A$ интерполяционные полиномы Лагранжа $L_n(f_1, A)$ неограничены на $[-1,+1]$. (Интерполяция непригодна.)

b) Если $A$ удовлетворяет условиям (1) и (2) и имеет место $\alpha > \beta$, то для любой $f \in \mathrm{Lip}\,\alpha$ интерполяционные полиномы Лагранжа $L_n(f, A)$ соответствующие $A$ стремятся к $f$ равномерно на $[-1, +1]$. (Интерполяция пригодна.)

c) Если $\alpha > \frac{\beta}{\beta+2}$, то уже существует матрица $A_0$, удовлетворяющая (1) и (2), для которой интерполяционные полиномы $L_n(f, A_0)$ равномерно сходятся к $f$, если только $f \in \mathrm{Lip}\,\alpha$.

d) Если $\alpha < \beta$, то уже существуют матрица $A_1$ удовлетворяющая (1) и (2), и функция $f_2 \in \mathrm{Lip}\,\alpha$ такие, что интерполяционные полиномы Лагранжа $L_n(f_2, A_1)$ неограничены в интервале $[-1, +1]$.

Аналогичные вопросы возникают также для разложений в ортогональный ряд, для механических квадратур и для других линейных операций.

# NOTES ON INTERPOLATION. I.
## (ON SOME INTERPOLATORICAL PROPERTIES OF THE ULTRASPHERICAL POLYNOMIALS)

By
J. SURÁNYI (Budapest) and P. TURÁN (Budapest), member of the Academy

*To* G. Szegő *on his 60th birthday*

**1.** Let be given $n$ points $x_1, \ldots, x_n$ with

$$-1 \leqq x_n < x_{n-1} < \cdots < x_1 \leqq 1. \tag{1.1}$$

It is a classical fact that there is exactly one polynomial $f(x)$ of degree $\leqq n-1$ such that $f(x)$ assumes at the place $x_\nu$ a prescribed value $y_\nu$. The explicite form of this polynomial $L_n(f)$ was given by Newton and Lagrange. The investigation of the convergence-properties of these polynomials, which began at the end of the last century by Runge and Borel, revealed a curious and unexpected situation with the results of G. Faber and S. Bernstein. We consider the system $A$ of points

$$-1 \leqq x_{nn} < x_{n-1,n} < \cdots < x_{2n} < x_{1n} \leqq 1 \qquad (n=1,2,\ldots) \tag{1.2}$$

and denote by $C$ the class of functions $f(x)$ continuous for $-1 \leqq x \leqq 1$. Then the above named authors proved that in whatever way we choose the system $A$, we can find an $f_0(x) \in C$ such that the polynomials $L_n(f_0)$ belonging to $A$ are unbounded in $[-1,+1]$.[1]

**2.** As Fejér[2] showed, the situation changes when we turn from the interpolatory polynomials $L_n(f)$ to another sort of interpolatory polynomials. Already Hermite[3] showed, that prescribing beside the points (1.1) the non-negative integers $m_1, \ldots, m_n$ and the real or complex numbers $y_{\mu\nu}$, there is a uniquely determined polynomial $f(x)$ of degree

$$\leqq m_1 + \cdots + m_n - 1 \tag{2.1}$$

[1] For this theory see e. g. И. П. Натансон, Конструктивная теория функций (Москва—Ленинград, 1949).

[2] L. Fejér, Über Interpolation, *Gött. Nachr.*, (1916), pp. 66—91.

[3] Ch. Hermite, Sur la formule d'interpolation de Lagrange, *Journ. für Math.*, 84 (1878), pp. 70—79.

such that

$$f^{(\nu)}(x_\mu) = y_{\mu\nu} \qquad (\mu = 1, 2, \ldots, n\ ;\ \nu = 0, 1, 2, \ldots, m_\mu - 1). \tag{2.2}$$

In the case

$$m_1 = m_2 = \cdots = m_n = 2 \tag{2.3}$$

FEJÉR gave a particularly simple form of these polynomials which enabled him in certain cases to perform the convergence-proofs in a very elegant manner. The new feature of the results is that for certain $A$-systems, choosing

$$y_{\mu 1} = 0 \qquad (\mu = 1, 2, \ldots, n),$$

the $H_n(f)$ Hermite—Fejér interpolatory polynomials of degree $\leqq 2n-1$ converge uniformly to $f(x)$ if only $f \in C$. In the more general cases of Hermite interpolation results are rather scarce, partly owing to the lack of suitably simple form of the interpolatory polynomials.

**3.** In all previously mentioned problems we prescribed at each point $x_\mu$ the value of the function and of some *consecutive* derivatives. In the most general case we consider a polynomial $f_m(x)$ of degree $\leqq (m_1 + \cdots + m_n - 1) = m$ such that for each $x_\mu$ we prescribe the value of some $m_\mu$ derivatives with given indices *(not necessarily consecutive ones!).* This general form of interpolation was treated as far as we know only in a paper of GEORGE BIRKHOFF.[4] His point of view — as the title of the paper shows — was so general that one cannot expect better formulae than those of HERMITE. One cannot even see from this paper the new feature of this general interpolation, viz. *for each $n$ by a suitable choice of the $x_\nu$'s and the indices of the prescribed derivatives the problem can be unsolvable or can have an infinity of solutions.* On this account, for the special case $n = 2$ a direct treatment was given by PÓLYA.[5] Hence the questions concerning this most general form of interpolation are roughly speaking the following ones:

1) Does there exist such an $f_m(x)$ at all? (Problem of existence.)

2) If there is such an $f_m(x)$, is it uniquely determined? (Problem of uniqueness.)

3) If $f_m(x)$ exists and is unique, how can it be represented in the most suitable form? (Problem of representation.)

4) Fixing for each $n$ the indices of the prescribed derivatives we choose the points (1.2), i. e. the system $A$ so that $f_m(x)$ should be uniquely determined, if the values of the prescribed derivatives are all 0 (if possible). Form-

[4] G. D. BIRKHOFF, General mean value and remainder theorems with applications to mechanical differentiation and quadrature, *Trans. Amer. Math. Soc.*, 7 (1906), pp. 107—136.

[5] G. PÓLYA, Bemerkung zur Interpolation und zur Näherungstheorie der Balkenbiegung, *Zeitschr. für angew. Math. und Mech.*, **11** (1931), pp. 445—449.

ing for an $f(x) \in C$ and each $n$ these interpolatory polynomials, does the sequence converge uniformly in $[-1, +1]$ to $f(x)$ or not? (Problem of convergence.)

**4.** In this sequence of papers we shall try to get orientation concerning this general sort of interpolation by considering the simplest case only, when $m = 2n-1$ and we are seeking a polynomial $f_{2n-1}(x)$ of degree $\leqq 2n-1$ where the values of the function and the *second* derivatives are prescribed at the points (1. 1), i. e.

(4. 1) $$\left.\begin{aligned} f_{2n-1}(x_\nu) &= y_\nu = \text{given}, \\ f''_{2n-1}(x_\nu) &= y^*_\nu = \text{given} \end{aligned}\right\} \qquad (\nu = 1, 2, \ldots, n).$$

The first question is how to choose the points (1. 1) in order to get the simplest proofs. It seems to us now that the best choice from this point of view are the zeros of the polynomial

(4. 2) $$\Pi_n(x) = (1-x^2) P'_{n-1}(x)$$

where $P_{n-1}(x)$ stands for the $(n-1)^{\text{th}}$ Legendre polynomial with the normalisation

(4. 3) $$P_{n-1}(1) = 1.$$

Denoting them by $\xi_{\mu\nu}$ we have[6]

(4. 4) $$-1 = \xi_{n,n} < \xi_{n-1,n} < \cdots < \xi_{2n} < \xi_{1n} = 1.$$

Forming these points for $n = 2, 3, \ldots$ and denoting this by $B$-system, in the second paper we shall consider the convergence behaviour of the corresponding interpolatory polynomials (4. 1) when $f(x) \in C$

(4. 5) $$y_\nu = f(\xi_{\nu n}), \quad y^*_\nu = 0 \qquad (\nu = 1, 2, \ldots, n;\ n = 2, 3, \ldots).$$

Here, in this first paper we shall consider the questions of existence and uniqueness only, but not exclusively for the $B$-matrix; already this will show how much simpler the proofs are in the case of the $B$-matrix.

**5.** Let

(5. 1) $$n = 2k+1$$

and

(5. 2) $$1 \geqq x_1 > x_2 > \cdots > x_k > x_{k+1} = 0 > x_{k+2} > \cdots > x_{2k+1} \geqq -1$$

with

(5. 3) $$x_j = -x_{2k+2-j} \qquad (j = k+2, \ldots, 2k+1).$$

We shall show the

THEOREM I. *If $n = 2k+1$ and the points $x_1, \ldots, x_n$ satisfy* (5. 2) *and* (5. 3), *there is in general no polynomial $f(x)$ of degree $\leqq 2n-1$ such that for*

[6] In this paper $n$ is fixed; thus the zeros of $\Pi_n(x)$ will be denoted simply by $\xi_\nu$ in the sequel.

*given $y_\nu$ and $y_\nu^*$*

$$f(x_\nu)=y_\nu,\quad f''(x_\nu)=y_\nu^* \qquad (\nu=1,2,\ldots,n). \tag{5.4}$$

*If there exists such a polynomial, then there is an infinity of them.*

Hence in the case of an odd number of distinct symmetrical points $x_\nu$ both the problem of existence and uniqueness have a negative solution. This means that in this case there is a universal linear relation between the values $\pi_{2n-1}(x_\nu)$ and $\pi_{2n-1}''(x_\nu)$ $(\nu=1,\ldots,n)$ of an arbitrary polynomial $\pi_{2n-1}(x)$ of degree $\leqq 2n-1$, the coefficients being *independent* of $\pi_{2n-1}(x)$. The proof will show that the same state of affairs holds if the second derivative is replaced by any derivative of fixed even order for $\nu=1,2,\ldots,n$.

**6.** The simple proof of Theorem I shall be given in **10.** The possibility of an infinity of solutions raises the question on the dimensionality of the number of them. Taking as points $x_\nu$ those defined in (4.4) we shall show directly in **11** *that in the case of infinitely many solutions in Theorem I for $n\geqq 3$ the general form of the solutions is*

$$f(x)=f_0(x)+cf_1(x), \tag{6.1}$$

*where $f_0(x)$ and $f_1(x)$ are fixed polynomials of degree $\leqq 2n-1$ and $c$ an arbitrary complex number.*

**7.** In the case $n=2k$ the situation changes. In **12** we shall show directly[7] the

THEOREM II. *If $n=2k$, then to prescribed values $y_\nu$ and $y_\nu^*$ there is a uniquely determined polynomial $f(x)$ of degree $\leqq 2n-1$ such that*

$$f(\xi_\nu)=y_\nu,\quad f''(\xi_\nu)=y_\nu^* \qquad (\nu=1,2,\ldots,n), \tag{7.1}$$

*if $\xi_\nu$ stands, as in (4.4), for the zeros of $\Pi_n(x)$.*

This means, of course, that in the case

$$y_\nu=y_\nu^*=0 \qquad (\nu=1,2,\ldots,n;\ n\text{ even}) \tag{7.2}$$

the only solution of (7.1) is $f(x)\equiv 0$. This result is curious for a certain reason. It is namely well known that $\Pi_n(x)$ satisfies the differential-equation (for each integer $n\geqq 2$)

$$(1-x^2)\Pi_n''(x)+n(n-1)\Pi_n(x)=0. \tag{7.3}$$

Thus for $n\geqq 4$

$$\Pi_n''(\xi_\nu)=0 \qquad (\nu=2,3,\ldots,(n-2),(n-1)).$$

Hence there is a non-trivial polynomial, even of $n^{\text{th}}$ degree, satisfying "almost" the requirements (7.2), i. e. except the two conditions

$$f''(\xi_1)=f''(\xi_n)=0. \tag{7.5}$$

[7] A verification of the further results of this paper in the frame of BIRKHOFF's approach seems to us to be a very difficult task.

Superficially thinking one would incline to conclude that there is a non-trivial polynomial of degree $(n+2)$ satisfying *all* conditions (7. 2). The Theorem II and the result in **11** show that this is not true; it follows that for an even $n$ no polynomial of degree $\leqq 2n-1$ exists with the property (7. 2) except $f\equiv 0$.

**8.** In **13** we shall investigate the existence- and uniqueness questions of the interpolation problem (4. 1) when the points $x_\nu$ are the zeros $\eta_{\nu n}^{(\lambda)}$ of the $n^{\text{th}}$ ultraspherical polynomial $P_n^{(\lambda)}(x)$, to the knowledge of which G. SZEGŐ contributed very much. Since $\lambda$ and $n$ are fixed we shall denote the zeros in question simply by $\eta_\nu$ $(\nu=1, 2, \ldots, n)$ and instead $P_n^{(\lambda)}(x)$ we shall write $Q_n(x)$ only. We consider only the case $\lambda \geqq -\frac{1}{2}$; since in the case $\lambda=-\frac{1}{2}$ $Q_n(x)$ is identical with $\Pi_n(x)$, we may suppose

$$\lambda>-\frac{1}{2}. \tag{8. 1}$$

The polynomial $Q_n(x)$ satisfies for all non-negative integer $n$'s the equation

$$(1-x^2)Q_n''(x)-(2\lambda+1)xQ_n'(x)+n(n+2\lambda)Q_n(x)=0; \tag{8. 2}$$

it is uniquely determined in the case $\lambda\neq 0$ by the normalisation

$$Q_n(1)=\binom{n+2\lambda-1}{n}, \tag{8. 3}$$

and for $\lambda=0$ by

$$Q_n(1)=1. \tag{8. 4}$$

It is well known that all the zeros are real, simple and lying in $-1<x<1$. Since for odd $n$ these zeros fulfil (5. 2) and (5. 3), we have only to confine ourselves in this § to the case

$$n=2k\geqq 4. \tag{8. 5}$$

We shall see that in this case the interpolation problem (4. 1) is generally uniquely determined. A particular behaviour is shown only in the case

$$\lambda+\frac{3}{2}=\text{odd integer} \qquad n\geqq \lambda+\frac{5}{2}, \tag{8. 6}$$

i. e. when the points $x_\nu$ are the zeros of some derivative of odd order of the Legendre polynomials. In this case neither existence nor uniqueness of the polynomials with (4. 1) holds good. More particular is the behaviour of these polynomials, if in (4. 1) we have

$$y_\nu=y_\nu^*=0 \qquad (\nu=1, 2, \ldots, n). \tag{8. 7}$$

In this case generally $f(x)\equiv 0$ is the only polynomial of degree $\leqq 2n-1$; but in the case (8. 6) there is a uniquely determined non-trivial polynomial

$f_0(x)$ of degree $\left(n+\lambda+\frac{1}{2}\right)$ such that for all polynomials $f_1(x)=cf_0(x)$ and only for these

$$f_1(\eta_\nu)=f_1''(\eta_\nu)=0 \qquad (\nu=1,2,\ldots,n) \tag{8.8}$$

holds with any numerical $c$. Thus we shall show the

THEOREM III. *The ultraspherical polynomials $P_n^{(\lambda)}(x)$, with $\lambda+\frac{1}{2}=$ even integer $>0$ and $n\geqq\lambda+\frac{5}{2}$, with the zeros $\eta_1,\ldots,\eta_n$, have the property that there exist non-trivial polynomials $f(x)$ of degree $\left(n+\lambda+\frac{1}{2}\right)$ such that $\eta_1,\ldots,\eta_n$ are zeros and at the same time places of inflexion of the polynomial $f(x)$.*

**9.** Before turning to the proofs we mention a corollary of Theorem I (similar conclusions could have been drawn from the other results of this paper, too). Let $n=2k+1$ and the numbers $x_\nu$ satisfy the condition (5.2)—(5.3); to a prescribed integrable $F(x)$ we consider the solution of the equation

$$f^{(2n)}(x)=F(x) \tag{9.1}$$

with the conditions

$$f(x_\nu)=y_\nu,\quad f''(x_\nu)=y_\nu^* \qquad (\nu=1,2,\ldots,n) \tag{9.2}$$

where $y_\nu$ and $y_\nu^*$ are prescribed numbers. If $g_0(x)$ is any solution of (9.1) (without requiring (9.2)), then $g(x)=f(x)-g_0(x)$ satisfies $g^{(2n)}(x)=0$, i. e.

$$f(x)=g_0(x)+\pi_{2n-1}(x)$$

where $\pi_{2n-1}(x)$ is a polynomial of degree $\leqq 2n-1$ such that

$$\pi_{2n-1}(x_\nu)=y_\nu-g_0(x_\nu)=\text{prescribed},$$
$$\pi_{2n-1}''(x_\nu)=y_\nu^*-g_0''(x_\nu)=\text{prescribed}.$$

But owing to Theorem I this is generally unsolvable. Hence we obtained the

COROLLARY.[8] *If $n$ is odd and the points $x_\nu$ are real and symmetrical to $x=0$, then the equation (9.1) is generally unsolvable under the restrictions (9.2).*

**10.** Now we turn to the proof of Theorem I. We decompose the polynomial

$$f(x)=\sum_{\nu=0}^{2n-1} c_\nu x^\nu$$

[8] A corollary of similar character is mentioned in BIRKHOFF's quoted paper, p. 114.

into an even and an odd part:

$$s(x)=\sum_{\nu=0}^{n-1} c_{2\nu}x^{2\nu}, \qquad t(x)=\sum_{\nu=0}^{n-1} c_{2\nu+1}x^{2\nu+1}.$$

The first part of the conditions (5. 4) will then give

$$\begin{aligned} f(x_\nu)&=s(x_\nu)+t(x_\nu)=y_\nu, \\ f(-x_\nu)&=s(x_\nu)-t(x_\nu)=y_{n-\nu+1} \qquad \left(\nu=1,2,\dots,\left[\frac{n}{2}\right]\right). \end{aligned} \tag{10. 1}$$

Since $s''(x)$ and $t''(x)$ are also even and odd polynomial, respectively, from the second part of the conditions follows

$$\begin{aligned} f''(x_\nu)&=s''(x_\nu)+t''(x_\nu)=y_\nu^*, \\ f''(-x_\nu)&=s''(x_\nu)-t''(x_\nu)=y_{n-\nu+1}^* \quad \left(\nu=1,2,\dots,\left[\frac{n}{2}\right]\right). \end{aligned} \tag{10. 2}$$

From (10. 1), (10. 2) we get for $t(x)$

$$t(x_\nu)=\frac{y_\nu-y_{n-\nu+1}}{2}=u_\nu,\ t''(x_\nu)=\frac{y_\nu^*-y_{n-\nu+1}^*}{2}=u_\nu^* \left(\nu=1,2,\dots,\left[\frac{n}{2}\right]\right), \tag{10. 3}$$

and of course

$$t(0)=t''(0)=0. \tag{10. 4}$$

Since the last conditions are automatically fulfilled, in the case of an odd $n$ we have $n-1$ linear equations for the $n$ coefficients of $t(x)$. Hence apart from special values $x_\nu$ the system (10. 3) has an infinity of solutions.

For $s(x)$ the equations (10. 1) and (10. 2) give

$$\begin{aligned} s(x_\nu)&=\frac{y_\nu+y_{n-\nu+1}}{2}=v_\nu, \\ s''(x_\nu)&=\frac{y_\nu^*+y_{n-\nu+1}^*}{2}=v_\nu^* \qquad \left(\nu=1,2,\dots,\left[\frac{n}{2}\right]\right). \end{aligned} \tag{10. 5}$$

Since in the case of an odd $n$, 0 occurs among the $x_\nu$'s, owing to (10. 4), we have the additional conditions

$$s(0)=y_{\left[\frac{n}{2}\right]+1}, \qquad s''(0)=y^*_{\left[\frac{n}{2}\right]+1}.$$

The last two equations determine $c_0$ and $c_2$. For the remaining $n-2$ coefficients the system (10.5) furnishes $n-1$ equations. This system becomes contradictory for appropriate values of the $v_\nu$'s and $v_\nu^*$'s.

Since the $u_\nu$'s and $v_\nu$'s resp. $u_\nu^*$'s and $v_\nu$'s also determine the values $y_\nu$ and $y_\nu^*$, these results prove Theorem I.

**11.** As told in **6** we shall consider with an odd $n$ in the case of $x_\nu=\xi_\nu$ (see (4. 4)) the general structure of the solutions. The assertion (6. 1) will be

obviously proved if we show that in our case *all* polynomials $f(x)$ of degree $\leqq 2n-1$ satisfying

$$f(\xi_\nu)=f''(\xi_\nu)=0 \qquad (\nu=1,2,\dots,n) \tag{11.1}$$

are

$$c\Pi_n(x)(P_{n-1}(x)-3) \tag{11.2}$$

with arbitrary numerical $c$. If $f(x)$ satisfies (11.1), the first part of the conditions implies

$$f(x)=\Pi_n(x)q_{n-1}(x)$$

where $q_{n-1}(x)$ is a polynomial of degree $\leqq n-1$. Then the second part of the conditions means

$$\Pi_n''(\xi_\nu)q_{n-1}(\xi_\nu)+2\Pi_n'(\xi_\nu)q_{n-1}'(\xi_\nu)=0 \qquad (\nu=1,2,\dots,(n-1),n). \tag{11.3}$$

But the equation (7.3) gives for $\nu=2,3,\dots,(n-2),(n-1)$

$$\Pi_n''(\xi_\nu)=0,$$

i. e. owing to the simplicity of the $\xi_\nu$'s we have

$$q_{n-1}'(\xi_\nu)=0 \qquad (\nu=2,3,\dots,(n-2),(n-1)).$$

But this means that $q_{n-1}'(x)$ has all its zeros common with

$$\frac{\Pi_n(x)}{1-x^2}=P_{n-1}'(x),$$

i. e.

$$q_{n-1}'(x)=cP_{n-1}'(x)$$

with a numerical $c$. Hence if $c\neq 0$

$$q_{n-1}(x)=c(P_{n-1}(x)+c_1)$$

with a numerical $c_1$, i. e.

$$f(x)=c\Pi_n(x)(P_{n-1}(x)+c_1). \tag{11.4}$$

To determine $c$ and $c_1$, we use the requirements $f''(\pm 1)=0$. As well known and easy to verify

$$P_{n-1}'(1)=\frac{n(n-1)}{2}=(-1)^n P_{n-1}'(-1) \tag{11.5}$$

and thus, using the known identity

$$\Pi_n(x)=-n(n-1)\int_{-1}^{x}P_{n-1}(t)dt, \tag{11.6}$$

we have

$$\Pi_n'(1)=-n(n-1)=(-1)^{n-1}\Pi_n'(-1), \tag{11.7}$$

$$\Pi_n''(1)=-\frac{n^2(n-1)^2}{2}=(-1)^n\Pi_n''(-1). \tag{11.8}$$

Hence from $f''(1)=0$ we have

$$0=c\{\Pi_n''(1)(1+c_1)+2\Pi_n'(1)P_{n-1}'(1)\}$$

and using (11. 5), (11. 7) and (11. 8)

$$c_1=-3,$$

(11. 9) $$f(x)=c\Pi_n(x)(P_{n-1}(x)-3).$$

Since $n$ is now odd, it is easy to verify with (11. 5), (11. 7) and (11. 8) that the polynomial (11. 9) satisfies also the condition $f''(-1)=0$. If $c=0$, then $q_{n-1}(x)=$ const., i. e. $f(x)$ would be $c_2\Pi_n(x)$ with a numerical $c_2$; but then $f''(1)\neq 0$.[9] Hence *only* the polynomials (11. 9) fulfil our requirement and thus (6. 1) is proved.

**12.** The proof of Theorem II runs parallel to the reasoning of **11.** This proves also in this case, $n$ being even, that if

(12. 1) $$f(\xi_\nu)=f''(\xi_\nu)=0,$$

then $f(x)$ has necessarily the form

(12. 2) $$f(x)=c\Pi_n(x)(P_{n-1}(x)-3)$$

with numerical $c$, without using the requirement $f''(-1)=0$. But since $n$ is now even, this condition together with (11. 5), (11. 7) and (11. 8) assures that $c=0$, i. e.

(12. 3) $$f(x)\equiv 0.$$

Now this means that writing out (12. 1) as a linear system, the determinant is $\neq 0$. Considering the general problem

(12. 4) $$f(\xi_\nu)=y_\nu,\quad f''(\xi_\nu)=y_\nu' \qquad (\nu=1,\dots,n)$$

this shows that the corresponding linear system is always uniquely soluble, i. e. Theorem II is proved.

**13.** Finally we turn, as told, to the case when the $x_\nu$'s are given by the numbers $\eta_\nu$ defined in **8.** Let first be

(13. 1) $$-\frac{1}{2}<\lambda\neq 0,\qquad n=2k.$$

As we have seen previously, everything depends on determination of all polynomials $g(x)$ of degree $\leqq 2n-1$ for which

(13. 2) $$g(\eta_\nu)=g''(\eta_\nu)=0 \qquad (\nu=1,2,\dots,n).$$

From the first part of the conditions we get

$$g(x)=Q_n(x)r_{n-1}(x)$$

[9] This last sentence holds obviously also in the case of an even $n$.

with a suitable polynomial $r_{n-1}(x)$ of degree $\leqq n-1$. The second part gives

(13. 3) $$Q_n''(\eta_\nu)r_{n-1}(\eta_\nu)+2Q_n'(\eta_\nu)r_{n-1}'(\eta_\nu)=0 \qquad (\nu=1, 2, \ldots, n).$$

But from (8. 2) we have

$$(1-\eta_\nu^2)Q_n''(\eta_\nu)-(2\lambda+1)\eta_\nu Q_n'(\eta_\nu)=0;$$

i. e. from (13. 3), using also the fact that the zeros of $Q_n(x)$ are simple, i. e. $Q_n'(\eta_\nu)\neq 0$, we get

(13. 4) $$\left(\lambda+\frac{1}{2}\right)\eta_\nu r_{n-1}(\eta_\nu)+(1-\eta_\nu^2)r_{n-1}'(\eta_\nu)=0.$$

Since the polynomial

$$(1-x^2)r_{n-1}'(x)+\left(\lambda+\frac{1}{2}\right)x r_{n-1}(x)$$

is of degree $\leqq n$ and by (13. 4) all its zeros are the same as those of $Q_n(x)$, we obtain

(13. 5) $$(1-x^2)r_{n-1}'(x)+\left(\lambda+\frac{1}{2}\right)x r_{n-1}(x)=cQ_n(x)$$

with numerical $c$.

**14.** We have to investigate whether or not the equation (13. 5) has a polynomial solution of degree $\leqq n-1$ ($n$ even). It would be easy to give the solution in the form of an integral, but it would not be easy to decide on this form whether or not the solution is a polynomial. Instead of it, we shall use the idea to try to solve the equation (13. 5) by

(14. 1) $$r_{n-1}(x)=\sum_{\nu=0}^{n-1} c_\nu Q_\nu(x).$$

We shall use the identities[10] for $\nu\geqq 1$

(14. 2) $$(1-x^2)Q_\nu'(x)= \\ =\frac{1}{2(\nu+\lambda)}\{(\nu+2\lambda-1)(\nu+2\lambda)Q_{\nu-1}(x)-\nu(\nu+1)Q_{\nu+1}(x)\},$$

(14. 3) $$xQ_\nu(x)=\frac{1}{2(\nu+\lambda)}\{(\nu+1)Q_{\nu+1}(x)+(\nu+2\lambda-1)Q_{\nu-1}(x)\}.$$

Substituting (14. 1) into (13. 5) we obtain

$$cQ_n(x)=\sum_{\nu=1}^{n-1} c_\nu(1-x^2)Q_\nu'(x)+\sum_{\nu=0}^{n-1} c_\nu\left(\lambda+\frac{1}{2}\right)xQ_\nu(x)$$

[10] See G. Szegő, *Orthogonal polynomials* (New York, 1939), pp. 83 and 82.

and using (14. 2), (14. 3)

$$c\,Q_n(x)=\sum_{\nu=1}^{n-1}\frac{c_\nu}{2(\nu+\lambda)}\{(\nu+2\lambda-1)(\nu+2\lambda)Q_{\nu-1}(x)-\nu(\nu+1)Q_{\nu+1}(x)\}+$$

$$+c_0\frac{\lambda+\frac{1}{2}}{2\lambda}Q_1(x)+\sum_{\nu=1}^{n-1}c_\nu\frac{\lambda+\frac{1}{2}}{2(\nu+\lambda)}\{(\nu+1)Q_{\nu+1}(x)+(\nu+2\lambda-1)Q_{\nu-1}(x)\}=$$

$$=c_0\frac{2\lambda+1}{4\lambda}Q_1(x)+\sum_{\nu=1}^{n-1}\frac{c_\nu}{2(\nu+\lambda)}(\nu+1)\left(\lambda+\frac{1}{2}-\nu\right)Q_{\nu+1}(x)+ \tag{14. 4}$$

$$+\sum_{\nu=1}^{n-1}\frac{c_\nu}{2(\nu+\lambda)}(\nu+2\lambda-1)\left(\nu+3\lambda+\frac{1}{2}\right)Q_{\nu-1}(x)=c_0\frac{2\lambda+1}{4\lambda}Q_1(x)+$$

$$+\sum_{\nu=2}^{n}\frac{c_{\nu-1}}{2(\nu+\lambda-1)}\nu\left(\lambda+\frac{3}{2}-\nu\right)Q_\nu(x)+$$

$$+\sum_{\nu=0}^{n-2}\frac{c_{\nu+1}}{2(\nu+\lambda+1)}(\nu+2\lambda)\left(\nu+3\lambda+\frac{3}{2}\right)Q_\nu(x).$$

We have to compare the coefficients of $Q_\nu(x)$ in (14. 4). Comparing the coefficients of $Q_{n-1}(x)$ we find for $n\geqq 4$

$$0=\frac{c_{n-2}}{2(n+\lambda-2)}(n-1)\left(\lambda+\frac{5}{2}-n\right),$$

i. e. if $n\geqq 4$ and $n\neq\lambda+\frac{5}{2}$

$$c_{n-2}=0. \tag{14. 5}$$

Comparing the coefficients of $Q_0(x)$ in (14. 4) we obtain

$$0=\frac{c_1}{\lambda+1}\lambda\left(3\lambda+\frac{3}{2}\right),$$

i. e. since now $\lambda\neq 0$

$$c_1=0. \tag{14. 6}$$

If $n\geqq 4$ and $2\leqq\nu\leqq n-2$, the comparison of the coefficients of $Q_\nu(x)$ in (14. 4) gives

$$0=\frac{c_{\nu-1}}{\nu+\lambda-1}\nu\left(\lambda+\frac{3}{2}-\nu\right)+\frac{c_{\nu+1}}{\nu+\lambda+1}(\nu+2\lambda)\left(\nu+3\lambda+\frac{3}{2}\right) \tag{14. 7}$$

and considering the coefficient of $Q_1(x)$ in (14. 4) we see that (14. 7) holds also for $\nu=1$; thus (14. 7) holds for

$$n\geqq 4,\quad 1\leqq\nu\leqq n-2. \tag{14. 8}$$

Evidently $c_{\nu+1}$ can be expressed always by $c_{\nu-1}$ from (14. 7); thus starting from (14. 6) we obtain

$$c_1=c_3=c_5=\cdots=c_{n-1}=0. \tag{14. 9}$$

Investigating $c_0, c_2, c_4, \ldots, c_{n-2}$ we consider first the case when $\lambda$ is not only restricted by (13. 1) but also by

$$\lambda + \frac{3}{2} \neq \text{odd integer.} \tag{14. 10}$$

Then $c_{\nu-1}$ can be expressed for odd $\nu$'s by $c_{\nu+1}$ from (14. 7) and hence starting from (14. 5) we obtain

$$c_{n-2} = c_{n-4} = \cdots = c_2 = c_0 = 0,$$

i. e. $r_{n-1}(x) \equiv 0$. Since (14. 10) implies that $n \neq \lambda + \frac{5}{2}$, and the case $\lambda = 0$ can be settled with reference to the continuity in $\lambda$, we obtained that *if* $n \geqq 4$ *even and*

$$\lambda + \frac{3}{2} \neq \textit{odd integer}, \quad \lambda > -\frac{1}{2}, \tag{14. 11}$$

*then we can always find a uniquely determined* $g(x)$ *polynomial of degree* $\leqq 2n-1$ *so that*

$$g(\eta_{\nu n}^{(\lambda)}) = \textit{prescribed}, \quad g''(\eta_{\nu n}^{(\lambda)}) = \textit{prescribed} \qquad (\nu = 1, 2, \ldots, n). \tag{14. 12}$$

**15.** If

$$\lambda > -\frac{1}{2}, \quad \lambda + \frac{3}{2} = \text{odd integer} \ (\geqq 3), \quad n = \text{even} \geqq \lambda + \frac{9}{2} \ (\geqq 6), \tag{15. 1}$$

the situation changes, since $c_{\nu-1}$ can be expressed by $c_{\nu+1}$ from (14. 7) only for $\nu > \lambda + \frac{3}{2}$, i. e. only

$$c_{n-2} = c_{n-4} = \cdots = c_{\lambda+\frac{5}{2}} = 0 \tag{15. 2}$$

remains true and $c_{\lambda+\frac{1}{2}}$ is indetermined. Giving any value to $c_{\lambda+\frac{1}{2}}$ (14. 7) is again applicable with

$$\nu = 1, 2, \ldots, \lambda - \frac{1}{2},$$

i. e. the values of

$$c_0, c_2, c_4, \ldots, c_{\lambda-\frac{3}{2}}$$

are with $c_{\lambda+\frac{1}{2}} \neq 0$ uniquely determined and $\neq 0$. Hence in this case all $g(x)$ polynomials of degree $\leqq 2n-1$ satisfying (14. 12) are of the form

$$g(x) = ck(x) \tag{15. 3}$$

where $k(x)$ is a uniquely determined polynomial of the exact degree $\left(n+\lambda+\frac{1}{2}\right)$. If

$$\lambda > -\frac{1}{2}, \quad \lambda+\frac{3}{2}=\text{odd integer}, \quad n=\lambda+\frac{5}{2}(\geqq 4), \tag{15. 4}$$

then (14. 5) ceases to be true, but

$$c_1 = c_3 = \cdots = c_{\lambda+\frac{3}{2}} = 0$$

holds. $c_{n-2}$ is now arbitrary, fixing its value, (14. 7) is applicable again with $\nu = 1, 2, \ldots, \left(\lambda - \frac{1}{2}\right)$, i. e. again $c_0, c_2, \ldots, c_{n-4} = c_{\lambda-\frac{3}{2}}$ are uniquely determined. Hence (15. 3) holds and the degree of $k(x)$ is $\lambda+\frac{1}{2}+n$ also in the case (15. 4). If the restrictions are

$$\lambda > -\frac{1}{2}, \quad \lambda+\frac{3}{2}=\text{odd integer } (\geqq 3), \quad 4 \leqq n = \text{even} \leqq \lambda+\frac{1}{2},$$

then repeating the same argument, it follows again $g(x) \equiv 0$. Thus we obtained that *if*

$$\lambda+\frac{3}{2}=\textit{odd integer } (\geqq 3), \tag{15. 5}$$

*then in the case* $4 \leqq n = \textit{even} \leqq \lambda+\frac{1}{2}$ *the only solution of the problem* (14. 12) *is* $g(x) \equiv 0$, *but for* $n \geqq \lambda+\frac{5}{2}$ *there is an infinity of solutions with the general form*

$$g(x) = ck(x),$$

*where* $c$ *is an arbitrary constant and* $k(x)$ *a uniquely determined polynomial of exact degree* $\left(n+\lambda+\frac{1}{2}\right)$.

*(Received 22 January 1955)*

# ЗАМЕЧАНИЕ ОБ ИНТЕРПОЛЯЦИИ. I. (О НЕКОТОРЫХ ИНТЕРПОЛЯЦИОННЫХ СВОЙСТВАХ УЛЬТРАСФЕРИЧЕСКИХ МНОГОЧЛЕНОВ)

Я. Шураньи и П. Туран (Будапешт)

(Резюме)

Авторы занимаются определением многочленов степени, не превышающей натурального числа $2n-1$, для которых значение функции и второй производной заданы на различных местах. Оказывается, что если $n$ — число нечетное и основные точки расположены симметрично относительно $O$, то задача никогда не является единственно определенной. При четном $n$ рассматривается случай, когда основные точки являются корнями $n$-ого ультрасферического многочлена, соответствующего некоторому знамению параметра $\lambda$; под ультрасферическим многочленом степени $n$ мы понимаем решение в виде многочлена степени $n$ уравнения (8.2). В этих случаях задача вообще единственно определена. Неопределенность возникает лишь если $\lambda+\frac{3}{2}$ есть нечетное целое число ($\geqq 3$). Особенно удобны для целей этих исследований основные точки, являющиеся корнями многочлена

$$\Pi_n(x)=\int_{-1}^{x} P_{n-1}(t)\,dt,$$

где $P_{n-1}(t)$ — многочлен Лежандра степени $n-1$. Мы вернемся к исследованию сходимости интерполяционного процесса вышеописанного типа, если основные точки корни многочлена $\Pi_n(x)$, в другом работе.

# ON PROCESSES OF HAPPENINGS GENERATED BY MEANS OF A POISSON PROCESS

By

LAJOS TAKÁCS (Budapest)

(*Presented by* A. RÉNYI)

## § 1. Introduction

In the second part of our paper [8] we have discussed the following problem:

Let us consider $m$ mutually independent stochastic processes, each defined as follows: let us consider a Poisson process in the time interval $0 \leqq u < \infty$, in which the density of the occurrence of events is $p$; some of these events produce a happening, the duration of which is a random variable. The first event taking place in the time interval $(0 \leqq u < \infty)$ produces a happening having the duration $\chi_1$, the first of the events taking place after the end of first happening produces a happening of duration $\chi_2$, etc. We suppose that the variables $\chi_k$ $(k=1, 2, 3, \ldots)$ are independent and equidistributed, with distribution function $P(\chi_k \leqq x) = H(x)$, further that $\chi_k$ can assume positive values only, that is $H(0)=0$ and that the expectation $M(\chi_k)=\alpha$ is finite. The secondary stochastic process obtained in this manner from the Poisson process is called *the process of happenings*. Let us mention that such processes occur in practice, for instance, in connection with the counting of particles by means of a Geiger—Müller counter.

In [8] it has been shown that the probability of a happening starting between the moments $u$ and $u+\Delta u$ is[1] $f(u)\Delta u + o(\Delta u)$ where $f(u)$ satisfies the following integral equation:

$$f(u)=p-p\int_0^u f(u-x)[1-H(x)]\,dx \tag{1}$$

which has a uniquely defined continuous solution $f(u)$. If the Laplace—Stieltjes

[1] By $o(\Delta u)$ we denote a function for which $\lim\limits_{\Delta u\to 0} \dfrac{o(\Delta u)}{\Delta u}=0$.

transform of $H(x)$ is

$$\psi(s)=\int_0^\infty e^{-sx}\,dH(x) \qquad (\Re(s)\geqq 0), \tag{2}$$

the Laplace transform of $f(u)$ will be

$$\int_0^\infty e^{-su} f(u)\,du=\frac{p}{s+p-p\psi(s)} \tag{3}$$

which converges for $\Re(s)>0$; relying upon this $f(u)$ can be uniquely determined.

The probability of a happening going on at the moment $u$ is

$$F(u)=1-\frac{1}{p}f(u). \tag{4}$$

It has been shown that the following limit exists:

$$f^*=\lim_{u\to\infty} f(u)=\frac{p}{1+p\alpha} \tag{5}$$

and consequently also

$$F^*=\lim_{u\to\infty} F(u)=\frac{p\alpha}{1+p\alpha}. \tag{6}$$

Let us now consider simultaneously $m$ processes of happenings of the type defined above. We shall say that at the moment $u$ the system is in the state $E_k$ $(k=0,1,\ldots,m)$ if at this moment in $m-k$ processes no happening is taking place, and in $k$ processes a happening is going on. In [8] the system we just referred to was examined first in the time interval $(0,t)$ and following that in any time interval having the length $t$, in case the process is going on since an infinite time. We call this latter case the stationary case. The expected numbers $m_{mk}(t)$ and $m^*_{mk}(t)$ of the transitions $E_{k-1}\to E_k$ during the time interval $(0,t)$ or, in the stationary case, during a time interval of the length $t$, respectively, have been determined. We say that at the moment $u$ an (at least $k$-fold) coincidence takes place if the system is in one of the states $E_j$ $(j=k,k+1,\ldots,m)$. We have determined $\tau_{mk}(t)$ and $\tau^*_{mk}(t)$, the mean total length of duration of the coincidences falling into the time intervall $(0,t)$ in the non-stationary case and in the stationary case, respectively.

We remark that the results concerning the stationary processes are to be understood that we determine the probabilities and expectations for the time interval $(u,u+t)$ and succeedingly we pass to the limit $u\to\infty$.

In the present paper we choose the special case $k=1$ as subject of a more detailed study. We shall determine in the non-stationary case the distribution of the number of transitions $E_0\to E_1$ in the time interval $(0,t)$; let us denote the probability of the event that this number is $\leqq n$ by $W_1(t,n)$ $(n=0,1,2,\ldots)$. Further, we determine the distribution of the total length of

time in the time interval $(0, t)$, while the system is in the states $E_j$ $(j=1,2,\ldots,m)$; let us denote this distribution function by $\Omega_1(t, z)$. These functions will be determined also for the stationary case and will be denoted by $W_1^*(t, n)$ and $\Omega_1^*(t, z)$, respectively.

Let us denote in succession the distances between a transition $E_0 \to E_1$ and the immediately following transition $E_1 \to E_0$ by the random variables $\xi_1, \xi_2, \ldots$. It is clear that $\xi_1, \xi_2, \ldots$ are independent random variables with the same distribution function $P(\xi_k \leqq x) = D(x)$ and expectation

(7) $$\vartheta = \int_0^\infty [1-D(x)]\,dx.$$

Let us denote in the case $k=1$ the above-mentioned mean values by $m_1(t), m_1^*(t), \tau_1(t)$ and $\tau_1^*(t)$, respectively. Relying on the results of [8] we have the following equations:

(8) $$m_1(t) = \sum_{n=0}^{\infty} [1-W_1(t, n)] = mp\int_0^t [1-F(u)]^m\,du,$$

(9) $$m_1^*(t) = \sum_{n=0}^{\infty} [1-W_1^*(t, n)] = mp\left(\frac{1}{1+p\alpha}\right)^m t$$

and

(10) $$\tau_1(t) = \int_0^t [1-\Omega_1(t, z)]\,dz = \int_0^t [1-(1-F(u))^m]\,du,$$

(11) $$\tau_1^*(t) = \int_0^t [1-\Omega_1^*(t, z)]\,dz = \left[1-\left(\frac{1}{1+p\alpha}\right)^m\right] t.$$

By comparison of (8) and (10), and of (9) and (11) it results

(12) $$m_1(t) + mp\tau_1(t) = mpt$$

and

(13) $$m_1^*(t) + mp\tau_1^*(t) = mpt.$$

The expectation $\vartheta$ can be determined using the renewal theory in the following manner:

(14) $$\vartheta = \lim_{t\to\infty} \frac{\tau_1^*(t)}{m_1^*(t)} = \frac{(1+p\alpha)^m - 1}{mp}.$$

In the first part of our paper we shall discuss the determination of the above-mentioned probabilities. The method employed is, as far as we know, new and at the same time very simple, and seems to be applied successfully in the solution of many problems of similar type. In the second part of our paper we shall discuss an apparently different problem by applying the same method. As we shall see, however, this problem is closely connected with the previous one, notably, it constitutes a limiting case of the previous problem.

## § 2. Preliminaries

In this section we prove some auxiliary theorems which will be needed in the following discussion.

Let us consider a stochastic process in the time interval $0 \leqq u < \infty$ describing the state of a system, with the only possible states $A$ and $B$. Let the system be at time $u=0$ in state $A$ and we suppose that the system passes in a random manner from state $A$ to $B$, and conversely. Let the durations of the intervals in which the system is staying in the states $A$ or $B$ be independent random variables. Let the successive durations, while the system is staying in state $A$, be denoted by the random variables $\eta_0, \eta_1, \eta_2, \ldots$. Let us suppose that the $\eta_k$'s have the same distribution function

$$C(x) = \begin{cases} 1-e^{-px} & \text{if } x \geqq 0, \\ 0 & \text{if } x < 0. \end{cases} \tag{15}$$

Further, we suppose that the successive durations of the system staying in state $B$, which we denote by $\xi_1, \xi_2, \xi_3, \ldots$, have the same distribution function $D(x)$ for which $D(0)=0$. Further, we suppose that the variables $\eta_0, \eta_1, \eta_2, \ldots; \xi_1, \xi_2, \ldots$ are mutually independent.

This definition determines the structure of the process uniquely. The regeneration points of the process are formed by points in which the system is in state $A$. Namely in such time points the probability that the following transition $A \rightarrow B$ occurs within the next time interval of length $x$ is $C(x)$, independently of the past of the system.

It is clear that the length of time between consecutive transitions $A \rightarrow B$, which we denote by $\zeta_1, \zeta_2, \zeta_3, \ldots$, are also equally distributed independent random variables with the same distribution function

$$G(x) = \int_0^x [1-e^{-p(x-y)}]\, dD(y), \tag{16}$$

since $\zeta_n = \eta_n + \xi_n$.

Let us denote by $W(t, n)$ the probability that the number of transition $A \rightarrow B$ occurring in the time interval $(0 \leqq u < t)$ is $\leqq n$, and by $W^*(t, n)$ in the stationary case. We shall determine these distribution functions as well as their mean values and moments.

We shall determine in the non-stationary case the distribution function $\Omega(t, z)$ of the total length of the duration within the time interval $(0, t)$, while the system is in state $B$. We shall determine the corresponding distribution function also in the stationary case and denote it by $\Omega^*(t, z)$. Further, we shall also determine the averages of these probability distributions.

## Non-stationary case

*The determination of* $W(t, n)$. Let us denote the instants of the transitions $A \to B$ by $u_1, u_2, \ldots, u_n, \ldots$. Then the probability of a number $\leqq n$ of transition $A \to B$ occurring in the time interval $(0 \leqq u \leqq t)$ is equal to the probability of the $n+1$-st transition occurring after the time $t$, that is to say

(17) $$W(t, n) = P(t < u_{n+1}) = 1 - P(u_{n+1} \leqq t) = 1 - C(t) * G_n(t),$$

where $G_n(t)$ denotes the $n$-fold convolution of the distribution function $G(t)$ which can be determined starting from $G_1(t) = G(t)$ by the following recursive formula:

(18) $$G_n(t) = \int_0^t G_{n-1}(t-x)\, d\, G(x) \qquad (n = 2, 3, \ldots).$$

Namely, $u_{n+1} = \eta_0 + \zeta_1 + \zeta_2 + \cdots + \zeta_n$ is the sum of $n+1$ independent random variables. We have $P(\eta_0 \leqq x) = C(x)$ and the distribution function

$$P(\zeta_1 + \zeta_2 + \cdots + \zeta_n \leqq x) = G_n(x)$$

satisfies to the recursive formula (18).

If the Laplace—Stieltjes transform of $G(x)$ is denoted by $\varphi(s)$, i. e.

(19) $$\varphi(s) = \int_0^\infty e^{-sx}\, d\, G(x),$$

if follows from (18) that the Laplace—Stieltjes transform of $\zeta_1 + \zeta_2 + \cdots + \zeta_n$ is

(20) $$\int_0^\infty e^{-sx}\, d\, G_n(x) = [\varphi(s)]^n.$$

Taking into consideration that

(21) $$\int_0^\infty e^{-sr}\, d\, C(x) = \frac{p}{p+s},$$

it follows from (17) that the Laplace transform of $W(t, n)$ will be

(22) $$\int_0^\infty e^{-st}\, W(t, n)\, dt = \frac{1 - [\varphi(s)]^{n+1}}{s}$$

whence $W(t, n)$ can be determined uniquely.

*Determination of the mean value.* If the expected number of transitions $A \to B$ occurring in the time interval $(0, t)$ is finite, what clearly holds in our case, we have

(23) $$m(t) = \sum_{n=0}^\infty [1 - W(t, n)] = \sum_{n=0}^\infty [C(t) * G_n(t)].$$

The Laplace—Stieltjes transform of (23) is

$$\int_0^\infty e^{-st}\,dm(t)=\frac{p}{p+s}\,\frac{1}{1-\varphi(s)}\,. \tag{24}$$

It is well known from the renewal theory (W. FELLER [3]) that the non-decreasing average function $m(t)$ can be determined uniquely from (24).

*Determination of the moments.* If the $r$-th moment of $W(t,n)$ is finite, it can be determined as follows:

$$m_r(t)=\sum_{n=0}^{\infty}[(n+1)^r-n^r]\,[1-W(t,n)]. \tag{25}$$

The Laplace—Stieltjes transform of this moment is

$$\int_0^\infty e^{-st}\,d\,m_r(t)=\frac{p}{p+s}\sum_{n=0}^{\infty}[(n+1)^r-n^r]\,[\varphi(s)]^n= \tag{26}$$

$$=\frac{p}{p+s}\sum_{j=1}^{r}\mathfrak{S}_r^j\,\frac{j!\,[\varphi(s)]^{j-1}}{[1-\varphi(s)]^j}\qquad(\Re(s)>0)$$

wherein $\mathfrak{S}_r^j$ denotes the Stirling numbers of the second kind. Here we have used the well-known relation

$$n^r=\sum_{j=1}^{r}\mathfrak{S}_r^j\,j!\binom{n}{j}. \tag{27}$$

By (26) $m_r(t)$ can be determined uniquely.

## Stationary case

*The determination of* $W^*(t,n)$. If the examined process has already been going on for an infinitely long time and we raise the question: what is the probability that in a time interval of length $t$ a number $\leqq n$ of transition $A\to B$ are taking place, it is necessary to know what is the situation at the starting point of the interval. The probability of the next transition occurring within a period of time $\leqq x$, counted from a given starting moment, is

$$G^*(x)=\frac{1}{\mu}\int_0^x[1-G(y)]\,dy \tag{28}$$

wherein $\mu$ denotes the average of $G(x)$, that is

$$\mu=\int_0^\infty[1-G(x)]\,dx. \tag{29}$$

This can be proved in the following manner. Let us denote by $\zeta_u$ the distance between the time point $u$ and the next transition $A \to B$. For the process $\zeta_u$ a theorem due to J. L. Doob [2] (Theorem 12) is applicable which states that under the present conditions there exists the limiting distribution $\lim_{u \to \infty} P(\zeta_u \leqq x) = G^*(x)$ where $G^*(x)$ is given by (28).

If the distribution function $G^*(x)$ is known, $W^*(t, n)$ can be determined in the same way as $W(t, n)$, only we must replace $\eta_0$ by such a random variable which has the distribution function $G^*(x)$. We can write

$$W^*(t, n) = 1 - G^*(t) * G_n(t). \tag{30}$$

As in our case

$$\int_0^\infty e^{-sx}\, d\, G^*(x) = \frac{1-\varphi(s)}{\mu s}, \tag{31}$$

we have by (20)

$$\int_0^\infty e^{-st}\, W^*(t, n)\, dt = \frac{1}{s} - \frac{[1-\varphi(s)]\,[\varphi(s)]^n}{\mu s^2} \qquad (\Re(s) > 0) \tag{32}$$

whence $W^*(t, n)$ can be determined uniquely.

*Determination of the mean value.* In the stationary case the expected number of events occurring during the period of time $t$ is

$$m^*(t) = \sum_{n=0}^{\infty} [1 - W^*(t, n)] = \frac{t}{\mu}, \tag{33}$$

what results by simple calculation, as it might have been expected in advance.

*Determination of the moments.* In case the moments of higher order are finite, they can be determined in a manner similar to what has been outlined in what precedes. The $r$-th moment is

$$m_r^*(t) = \sum_{n=0}^{\infty} [(n+1)^r - n^r]\,[1 - W^*(t, n)], \tag{34}$$

its Laplace—Stieltjes transform is similarly to (28),

$$\int_0^\infty e^{-st}\, d\, m_r^*(t) = \frac{1}{\mu s} \sum_{j=1}^{r} \mathfrak{S}_r^j \, \frac{j!\,[\varphi(s)]^{j-1}}{[1-\varphi(s)]^{j-1}} \qquad (\Re(s) > 0) \tag{35}$$

and hence $m_r^*(t)$ can be determined.

## Duration of staying in the state $B$

As we have seen the distribution function $G(x)$, of the length of time between two consecutive transitions $A \to B$, can be determined from (16). On the other hand, if we know $G(x)$, $D(x)$ can be determined from (16). As the Laplace—Stieltjes transform of $G(x)$ is $\varphi(s)$ and the Laplace—Stieltjes transform of $C(x)$ is $p/(p+s)$, it follows from (16)

$$\int_0^\infty e^{-sx}\, dD(x) = \frac{\varphi(s)(p+s)}{p} \qquad (\Re(s) \geqq 0) \tag{36}$$

whence the distribution function $D(x)$ can be determined uniquely. Let us denote the $n$-fold convolution of $D(x)$ by $D_n(x)$. This can be determined very simply by aid of (36) since according to that formula the Laplace—Stieltjes transform of $D_n(x)$ is

$$\int_0^\infty e^{-sx}\, dD_n(x) = \left[\frac{\varphi(s)(p+s)}{p}\right]^n \qquad (\Re(s) \geqq 0) \tag{37}$$

whence $D_n(x)$ can be determined. We put $D_0(x) = 0$ if $x < 0$, and $D_0(x) = 1$ if $x \geqq 0$.

In order to determine the probability distribution functions $\Omega(t, z)$ and $\Omega^*(t, z)$ we refer to our paper [8]: in the non-stationary case we get

$$\Omega(t, z) = \begin{cases} \sum_{n=0}^{\infty} e^{-p(t-z)} \dfrac{[p(t-z)]^n}{n!} D_n(z) & \text{if } 0 \leqq z \leqq t, \\ 1 & \text{if } z \geqq t \end{cases} \tag{38}$$

and in the stationary case

$$\Omega^*(t, z) = \begin{cases} \dfrac{1}{1+p\vartheta} \sum_{n=0}^{\infty} e^{-p(t-z)} \dfrac{[p(t-z)]^n}{n!} \left\{ D_n(z) + p\int_0^z [1-D(x)]\, D_n(z-x)\, dx \right\} & \text{if } 0 \leqq z < t, \\ 1 & \text{if } z \geqq t, \end{cases} \tag{39}$$

where $\vartheta$ denotes the mean value of the duration of a single staying of state $B$:

$$\vartheta = \int_0^\infty [1-D(x)]\, dx. \tag{40}$$

We remark that $\Omega^*(t, z)$ has a jump

$$\frac{p}{1+p\vartheta} \int_t^\infty [1-D(x)]\, dx \tag{41}$$

at $z=t$. The probabilities $\Omega(t,z)$ and $\Omega^*(t,z)$ can also be obtained in a simple manner by aid of Laplace transformation, notably, we would confine ourselves to the case in which $t-z=a$ ($a$ constant) and the Laplace transform with respect to $z$ of $\Omega(a+z,z)$ and $\Omega^*(a+z,z)$, are formed with the parameter $a$. In this case the following equation holds:

$$\int_0^\infty e^{-sz}\Omega(a+z,z)\,dz=\frac{1}{s}\sum_{n=0}^{\infty}e^{-pa}\frac{(pa)^n}{n!}\left[\frac{\varphi(s)(p+s)}{p}\right]^n=\frac{1}{s}e^{-a[p-(p+s)\varphi(s)]}, \tag{42}$$

and similarly

$$\int_0^\infty e^{-sz}\Omega^*(a+z,z)\,dz=\frac{(p+s)(1-\varphi(s))}{(1+p\vartheta)s^2}e^{-a[p-(p+s)\varphi(s)]}. \tag{43}$$

By inversion of the Laplace transform we obtain $\Omega(a+z,z)$ and $\Omega^*(a+z,z)$, respectively. From these the desired probability distribution functions $\Omega(t,z)$ and $\Omega^*(t,z)$ may be obtained by the substitution $a=t-z$.

The average duration of the system staying in state $B$ is in the non-stationary case

$$\tau(t)=\int_0^t[1-\Omega(t,z)]\,dz \tag{44}$$

and its Laplace—Stieltjes transform is

$$\int_0^\infty e^{-st}\,d\tau(t)=\frac{1}{s}-\frac{1}{p+s}\sum_{n=0}^{\infty}[\varphi(s)]^n=\frac{1}{s}-\frac{1}{(p+s)[1-\varphi(s)]}\quad(\Re(s)>0). \tag{45}$$

In the stationary case we have

$$\tau^*(t)=\int_0^t[1-\Omega^*(t,z)]\,dz=\frac{pt}{1+p\vartheta}, \tag{46}$$

as results by simple calculation.

We shall not deal with the question of determining the moments of higher order; in this respect we refer to [8].

It appears from our above discussion that if the distribution function $G(x)$ is known, it will be possible, by considering its Laplace—Stieltjes transform $\varphi(s)$, to determine all the probabilities in question. In what follows we shall discuss cases, in which the distribution function $G(x)$ is not known, but, indirectly, we are able to determine the average function $m(t)$. If $m(t)$ is known, $\varphi(s)$ can be determined according (24) by

$$\varphi(s)=\int_0^\infty e^{-sx}\,dG(x)=1-\frac{p}{p+s}\,\frac{1}{\int_0^\infty e^{-st}\,dm(t)}, \tag{47}$$

the mean value of $G(x)$ is evidently

$$\mu=\int_0^\infty x\,dG(x)=\lim_{t\to\infty}\frac{t}{m(t)}. \tag{48}$$

If we denote by $P(t)$ the probability that the system is at time $t$ in state $A$ and $\lim_{t\to\infty} P(t)=P$ exists, then relying our results [8] we have $\mu=1/pP$.

The main idea of our paper is the following: there exist cases in which the determination of $G(x)$ either can not be carried out or raises great difficulties, whilst at the same time $m(t)$ can be calculated in a simple manner, succeedingly $\varphi(s)$, the Laplace—Stieltjes transform of $G(x)$, is obtained very simply and thus also $G(x)$ itself can be determined.

It must be mentioned that in the next two sections we shall use the same notations as in the general case treated above.

## § 3. The solution of the problem

As easily seen, in the system consisting of $m$ mutually independent processes of happenings the moments in which the system is in state $E_0$ form the regeneration points of the process. The durations of the time intervals in which the system is in state $E_0$ are independent random variables with the same distribution function $C(x)=1-e^{-mpx}$. The lengths of successive time intervals in which the system is in state *non*-$E_0$ (that is to say in one of the states $E_1, E_2, \ldots, E_m$) are also independent random variables with the same unknown distribution function $D(x)$.

This is the same process as discussed in the last chapter. To the state $E_0$ corresponds, in the preceding chapter the state $A$, to the state *non*-$E_0$ the state $B$ and to the transition $E_0\to E_1$ the transition $A\to B$. Let us denote the common distribution functions of the distances between consecutive transitions $E_0\to E_1$ by $G(x)$.

We consider the Laplace—Stieltjes transform of $G(x)$

$$\varphi(s)=\int_0^\infty e^{-sx}\,dG(x) \qquad (\Re(s)\geq 0). \tag{49}$$

If this transform is known, it will be possible, using the results of the previous chapter, to determine the probability in question.

Thus $W_1(t,n)$, the probability of $\leq n$ transition $E_0\to E_1$ taking place in the interval $(0,t)$, can be determined uniquely by the following expression:

$$\int_0^\infty e^{-st}W_1(t,n)\,dt=\frac{1}{s}-\frac{mp}{mp+s}\,\frac{[\varphi(s)]^n}{s} \tag{50}$$

which converges for $\Re(s)>0$.

In the stationary case $W_1^*(t, n)$ can be determined from the transform

$$\int_0^\infty e^{-st} W_1^*(t, n)\,dt = \frac{1}{s} - \frac{[1-\varphi(s)]\,[\varphi(s)]^n}{\mu s^2} \qquad (\Re(s)>0). \tag{51}$$

Here $\mu$ denotes the average of $G(x)$.

In our case, in accordance with (36), the following equation holds:

$$\int_0^\infty e^{-sx}\,dD(x) = \frac{mp+s}{mp}\,\varphi(s). \tag{52}$$

Let the mean value of $D(x)$ be $\vartheta = \mu - 1\ mp$ and the $n$-fold convolution of $D(x)$ be $D_n(x)$; in this case, in accordance with (38), the following equation holds:

$$\Omega_1(t, z) = \sum_{n=0}^\infty e^{-mp(t-z)} \frac{[mp(t-z)]^n}{n!} D_n(z) \qquad (0 \leqq z \leqq t) \tag{53}$$

and by (39)

$$\begin{aligned} \Omega_1^*(t, z) &= \frac{1}{1+mp\vartheta} \sum_{n=0}^\infty e^{-mp(t-z)} \frac{[mp(t-z)]^n}{n!}\,\cdot \\ &\cdot \left\{ D_n(z) + mp\int_0^z [1-D(x)]\,D_n(z-x)\,dx \right\} \qquad (0 \leqq z < t). \end{aligned} \tag{54}$$

As can be seen, it is possible to answer all the questions if $\varphi(s)$ is known. This will be obtained by aid of (47) $m_1(t)$. According to (8), taking (4) into account,

$$m_1(t) = \frac{mp}{p^m} \int_0^t [f(u)]^m\,du \tag{55}$$

and the Laplace—Stieltjes transform of this function is

$$\int_0^\infty e^{-st}\,dm_1(t) = \frac{mp}{p^m} \int_0^\infty e^{-st}[f(t)]^m\,dt \qquad (\Re(s)>0). \tag{56}$$

Hence, according to (47), we obtain

$$\varphi(s) = 1 - \frac{p^m}{mp+s} \left\{ \int_0^\infty e^{-st}[f(t)]^m\,dt \right\}^{-1}. \tag{57}$$

Thus, $\varphi(s)$ having been determined, the probabilities in question can also be obtained in a simple manner. We would remark that with a given

$H(x)$ the solution $f(u)$ is furnished by (1). The mean value of $G(x)$ is

$$\mu=\int_0^\infty x\,dG(x)=\frac{(1+p\alpha)^m}{mp} \tag{58}$$

where $\alpha=M\{\chi_k\}$, i.e. the expectation of $\chi_k$.

We remark that, generally, our method can not be employed for studying the transitions $E_k\to E_{k+1}$ or $E_{k+1}\to E_k$ $(k\neq 0)$ because the periods elapsing between such transitions depend on the length of time during which the happenings going on at the starting moment endure. The only exception is the case in which the variables $\chi_k$ are exponentially distributed which may be treated easily by the same method. For the case $k=0$, however, the method described is suitable for the general solution of our problem.

## § 4. Processes of happenings generated by a Poisson process

For the application of the method mentioned above we shall give another example which can be discussed by aid of this method in a very simple manner, whereas otherwise its discussion would be very complicated, even in special cases. As will be seen, this example stands in close connection with the preceding one.

Let us consider a Poisson process of density $\lambda$ in the time interval $0\leqq u<\infty$. We suppose that each event produces a happening lasting for a period of time $\chi_k$ $(k=1,2,\ldots)$, where the $\chi_k$'s are independent random variables having the same distribution function $P(\chi_k\leqq x)=H(x)$. Let us suppose that $\chi_k>0$, that is $H(0)=0$. This process has been examined by A. RÉNYI [7], who proved that the number of happenings going on at a given moment has a Poisson distribution.

We shall say that the process is at the moment $u$ in the state $E_k$ if the number of happenings just going on is $k$. Let $W_1(t,n)$ denote the probability of the event that the number of transitions $E_0\to E_1$ in the time interval $(0,t)$ is $\leqq n$ and let us designe the same probability in the stationary case in an interval of the length $t$ by $W_1^*(t,n)$. The distribution function of the total length of time, while the process is in the state *non*-$E_0$ (that is one of the states $E_j$, $j=1,2,\ldots$) in the time interval $(0,t)$, will be denoted by $\Omega_1(t,z)$ and, correspondingly, in the stationary case by $\Omega_1^*(t,z)$.

As easily seen, those instants in which the system is in state $E_0$ form the Markov points of the process. The lengths of successive time intervals in which the system is in state $E_0$ are independent random variables with the common distribution function $C(x)=1-e^{-\lambda x}$. The durations in which the system is in state *non*-$E_0$ are also independent random variables with the common unknown distribution function $D(x)$. This process agrees with the

process discussed in the second chapter if we let correspond $A$ to $E_0$, $B$ to *non*-$E_0$ and the transitions $A \to B$ to the transitions $E_0 \to E_1$. Accordingly, the distances between consecutive transitions $E_0 \to E_1$ are independent random variables having the common distribution function $G(x) = C(x) * D(x)$. Let us denote the mean value of $G(x)$ by $\mu$. Let us consider the Laplace—Stieltjes transform of $G(x)$

(59) $$\varphi(s) = \int_0^\infty e^{-sx} dG(x).$$

Thus we obtain by means of (16)

(60) $$\int_0^\infty e^{-sx} dD(x) = \frac{\lambda + s}{\lambda} \varphi(s).$$

Let us denote the $n$-fold convolution of $D(x)$ by $D_n(x)$ and the average of $D(x)$ by $\vartheta = \mu - \frac{1}{\lambda}$.

Relying on the previous chapters,

(61) $$\int_0^\infty e^{-st} W_1(t, n) dt = \frac{1}{s} - \frac{\lambda}{\lambda + s} \frac{[\varphi(s)]^n}{s},$$

(62) $$\int_0^\infty e^{-st} W_1^*(t, n) dt = \frac{1}{s} - \frac{[1 - \varphi(s)][\varphi(s)]^n}{\mu s^2}$$

and

(63) $$\Omega_1(t, z) = \sum_{n=0}^\infty e^{-\lambda(t-z)} \frac{[\lambda(t-z)]^n}{n!} D_n(z) \qquad (0 \leqq z \leqq t),$$

(64) $$\Omega_1^*(t, z) = \begin{cases} \dfrac{1}{1 + \lambda\vartheta} \displaystyle\sum_{n=0}^\infty e^{-\lambda(t-z)} \frac{[\lambda(t-z)]^n}{n!} \cdot \\ \quad \cdot \left\{ D_n(z) + \lambda \displaystyle\int_0^z [1 - D(x)] D_n(z - x) dx \right\} & \text{if } 0 \leqq z < t, \\ 1 & \text{if } z \geqq t. \end{cases}$$

It appears that the determination of the probabilities in question has been reduced to the determination of $\varphi(s)$ and this, in turn, is indirectly obtained by determining the average function.

The above problem has been discussed by C. LEVERT and W. L. SCHEEN [6], L. KOSTEN [5] and W. FELLER [4] in connection with Geiger—Müller counters, whilst in connection with the distribution of random intervals along a straight line it has been discussed by C. DOMB [1]. These authors discussed only the case in which $\chi_k = \alpha$ (constant). LEVERT and SCHEEN substituted

the time interval of length $t$ by the periphery of a circle and calculated by means of complicated multi-dimensional integrals, for the problem thus modified, the corresponding probability $W_1(t, n)$ (this, of course, differs from ours); further, they determined the factorial moments of this probability. KOSTEN criticises the discussion of the problem by LEVERT and SCHEEN; for the stationary case he indicates the exact solution $W_1^*(t, n)$ and calculates its factorial moment. His method is based on the solution by means of the Heaviside operator calculus, of equilibrium equations composed of a system of partial differential equations and integral equations. W. FELLER reduces the problem to the theory of summation of random variables and so he obtains simpler results than the previous authors. DOMB determines the distribution function $\Omega_1^*(t, z)$, or at least its Laplace transform, by aid of recursive formulae which he solves by aid of Laplace transformation.

We shall show how simply the above results may be obtained with our method.

All that we require is the determination of $\varphi(s)$. This can be determined, according to (47), by aid of the average function. If we denote by $m_1(t)$ the expexted number of transitions $E_0 \to E_1$ occurring in time interval $(0, t)$, then $m_1(t+\Delta t)-m_1(t)$ is equal to the expected number of transitions $E_0 \to E_1$ occurring in the time interval $(t, t+\Delta t)$. Let us denote by $P_0(t)$ the probability that the system is in state $E_0$ at the moment $t$. We have

(65)
$$P_0(t)=\begin{cases} e^{-\lambda t} & \text{if} \quad 0 \leqq t \leqq \alpha, \\ e^{-\lambda\alpha} & \text{if} \quad t \geqq \alpha. \end{cases}$$

Namely, the system is at time $t$ in state $E_0$, if in the interval $(0, t)$ or $(t-\alpha, t)$, respectively, there does not occur any event.

We obtain

(66)
$$m_1(t+\Delta t)-m_1(t)=P_0(t)\lambda\Delta t+o(\Delta t).$$

If the system is at time $t$ in state $E_0$, then in the time interval $(t, t+\Delta t)$ there is one event occurring with the probability $\lambda\Delta t+o(\Delta t)$ and more than one with the probability $o(\Delta t)$. In other cases the expected number of the transitions $E_0 \to E_1$ in time interval $(t, t+\Delta t)$ is $o(\Delta t)$.

We have from (66)

(67)
$$m_1'(t)=\begin{cases} e^{-\lambda t}\lambda & \text{if} \quad 0 \leqq t \leqq \alpha, \\ e^{-\lambda\alpha}\lambda & \text{if} \quad t \geqq \alpha. \end{cases}$$

Hence

(68)
$$\int_0^\infty e^{-st}\,dm_1(t)=\frac{\lambda}{\lambda+s}\,\frac{s+\lambda e^{-(\lambda+s)\alpha}}{s}$$

and with the aid of (47) we obtain

(69)
$$\varphi(s)=\frac{\lambda e^{-\alpha(\lambda+s)}}{s+\lambda e^{-\alpha(\lambda+s)}}.$$

Knowing this the probabilities in question can be determined by means of the formulae (61)—(64).

After these preliminary remarks we are passing to the general case, where $H(x)$, the distribution function of $\chi_k$, is arbitrary. Considering the happenings beginning in the interval $(0, t)$, it appears that, as it has been shown by A. RÉNYI [7], the probability of the system being at the moment $t$ in the state $E_0$ is

$$P_0(t) = e^{-\lambda \int_0^t [1-H(x)]\,dx}. \tag{70}$$

This formula can be simply deduced as follows:

$$P_0(t) = e^{-\lambda t} + \lambda \int_0^t e^{-\lambda(t-x)} P_0(x) H(x)\,dx, \tag{71}$$

because the system may be in state $E_0$ at the time $t$ in one of both following ways: Either in the time interval $(0, t)$ does not occur any event, the probability whereof is $e^{-\lambda t}$; or in the time interval $(0, t)$ the first transition $E_0 \to E_1$ occurs at the time $t-x$ (the density function for which is $e^{-\lambda(t-x)}\lambda$) and the produced happening terminated not later than at the time $t$, the probability whereof is $H(x)$, and all the happenings beginning in the time interval $(t-x, t)$ finish before the moment $t$, the probability whereof is $P_0(x)$. Putting $g(t) = e^{\lambda t} P_0(t)$, we have

$$g(t) = 1 + \lambda \int_0^t g(x) H(x)\,dx \tag{72}$$

where $g(0) = 1$. By differentiation we obtain

$$g'(t) = \lambda g(t) H(t). \tag{73}$$

Hence

$$g(t) = e^{\lambda \int_0^t H(x)\,dx} \tag{74}$$

and finally

$$P_0(t) = e^{-\lambda \int_0^t [1-H(x)]\,dx}. \tag{75}$$

Now let us determine the expected number of transitions $E_0 \to E_1$ occurring in the time interval $(t, t+\Delta t)$. This yields the following expression:

$$m_1(t+\Delta t) - m_1(t) = P_0(t)\lambda\Delta t + o(\Delta t). \tag{76}$$

If the system is in state $E_0$ at time $t$, then in the time interval $(t, t+\Delta t)$ there occurs one event with probability $\lambda\Delta t + o(\Delta t)$ and more than one event with probability $o(\Delta t)$. In other cases the expected number of the transitions $E_0 \to E_1$ in time interval $(t, t+\Delta t)$ is $o(\Delta t)$.

From (76) we obtain

$$m_1'(t) = \lambda P_0(t) \tag{77}$$

and thus

$$\int_0^\infty e^{-st} dm_1(t) = \lambda \int_0^\infty e^{-st} P_0(t) dt = \lambda \int_0^\infty e^{-st-\lambda \int_0^t [1-H(x)]dx} dt, \tag{78}$$

whence, by (47), we obtain

$$\varphi(s) = 1 - \frac{1}{\lambda+s} \left\{ \int_0^\infty e^{-st-\int_0^t [1-H(x)]dx} dt \right\}^{-1}. \tag{79}$$

Knowing this, the probabilities in question can be calculated by aid of the formulae (61)—(64).

Now we have

$$\mu = \int_0^\infty x\,dG(x) = \frac{e^{\alpha\lambda}}{\lambda} \tag{80}$$

where $\alpha = M\{\chi_k\}$.

The process discussed in this last chapter stands in close connection with the system consisting of $m$ processes of happenings discussed earlier. Notably, this process constitutes the limiting case of the earlier discussed one, reached when $m \to \infty$, $p \to 0$ and $mp = \lambda$ (constant). To prove this, it is sufficient to show that, while passing to this limit the formula (57) goes over into the formula (79).

Using (1) it can be shown that the following equation holds:

$$f(t) = p - p^2 \int_0^t [1-H(x)]dx + o(p^2) \tag{81}$$

and thus we obtain

$$\left[\frac{f(t)}{p}\right]^m = [1 - p\int_0^t [1-H(x)]dx + o(p)]^m \to e^{-\lambda \int_0^t [1-H(x)]dx} \quad \text{as } m \to \infty. \tag{82}$$

By employing this latter formula (79) will follow from (57). The fact that the second process is the limiting case of the preceding one is intuitively also evident.

Example. Let $\chi_k = \alpha$, i. e. a constant, that is to say, put $H(x) = 0$ if $x < \alpha$ and $H(x) = 1$ if $x \geq \alpha$. In this case it follows according to (79)

$$\varphi(s) = \frac{\lambda e^{-\alpha(\lambda+s)}}{s + \lambda e^{-\alpha(\lambda+s)}} \tag{83}$$

in conformity with (69). Hence

$$G(x)=\sum_{j=1}^{\left[\frac{x}{\alpha}\right]}(-1)^{j-1}\frac{e^{-j\alpha\lambda}\lambda^j}{j!}(x-j\alpha)^j \tag{84}$$

and, according to (60), the Laplace—Stieltjes transform of $D(x)$ will be

$$\int_0^\infty e^{-sx}\,dD(x)=\frac{\lambda+s}{\lambda}\,\frac{\lambda e^{-\alpha(\lambda+s)}}{s+\lambda e^{-\alpha(\lambda+s)}} \tag{85}$$

whence

$$D(x)=\sum_{j=1}^{\left[\frac{x}{\alpha}\right]}(-1)^{j-1}e^{-j\alpha\lambda}\left[\frac{\lambda^j(x-j\alpha)^j}{j!}+\frac{\lambda^{j-1}(x-j\alpha)^{j-1}}{(j-1)!}\right] \tag{86}$$

and similarly

$$D_n(x)=\sum_{j=n}^{\left[\frac{x}{\alpha}\right]}(-1)^{j-n}\binom{j-1}{n-1}e^{-j\alpha\lambda}\left[\sum_{k=0}^{n}\binom{n}{k}\frac{\lambda^{j-k}(x-j\alpha)^{j-k}}{(j-k)!}\right]. \tag{87}$$

According to (61) we obtain

$$\begin{aligned}\int_0^\infty e^{-st}W_1(t,n)\,dt&=\frac{1}{s}-\frac{\lambda}{\lambda+s}\,\frac{[\varphi(s)]^n}{s}=\\&=\frac{1}{s}-\frac{\lambda}{\lambda+s}\sum_{j=n}^{\infty}(-1)^{j-n}\binom{j-1}{n-1}\frac{\lambda^j e^{-j\alpha(\lambda+s)}}{s^{j+1}}\end{aligned} \tag{88}$$

whence

$$W_1(t,n)=1+(-1)^n\sum_{j=n}^{\left[\frac{t}{\alpha}\right]}\binom{j-1}{n-1}e^{-j\alpha\lambda}\left[e^{-\lambda(t-j\alpha)}-\sum_{k=0}^{j}(-1)^k\frac{\lambda^k(t-j\alpha)^k}{k!}\right] \tag{89}$$

and according to (62)

$$\begin{aligned}\int_0^\infty e^{-st}W_1^*(t,n)\,dt&=\frac{1}{s}-\frac{1-\varphi(s)}{\mu s}\,\frac{[\varphi(s)]^n}{s}=\\&=\frac{1}{s}-\frac{1}{\mu}\sum_{j=n}^{\infty}(-1)^{j-n}\binom{j}{n}\frac{\lambda^j e^{-\alpha j(\lambda+s)}}{s^{j+2}}\end{aligned} \tag{90}$$

wherein $\mu$ denotes the average of $G(x)$, the value whereof is, by simple calculation,

$$\mu=\frac{e^{\alpha\lambda}}{\lambda}. \tag{91}$$

Hence

$$(92) \qquad W_1^*(t,n) = 1 - \sum_{j=n}^{\left[\frac{t}{\alpha}\right]} (-1)^{j-n} \binom{j}{n} \frac{e^{-\alpha\lambda(j+1)}\lambda^{j+1}(t-j\alpha)^{j+1}}{(j+1)!}$$

in conformity with the result of KOSTEN [5].

The distribution functions $\Omega_1(t,z)$ and $\Omega_1^*(t,z)$ may be written in accordance with (63) and (64) by aid of the expression (87) of $D_n(x)$. Now $\vartheta = (e^{\alpha\lambda} - 1)/\lambda$ is the average of $D(x)$. It is only an appearence that the expressions $\Omega_1(t,z)$ and $\Omega_1^*(t.z)$ contain an infinite number of terms, because $D_n(x) = 0$ if $n < x/\alpha$.

INSTITUTE FOR APPLIED MATHEMATICS
OF THE HUNGARIAN ACADEMY OF SCIENCES

*(Received 17 August 1952)*

## Bibliography

[1] C. DOMB, The problem of random intervals on a line, *Proc. Cambridge Phil. Soc.*, **43** (1947), pp. 329—341.
[2] J. L. DOOB, Renewal theory from the point of view of the theory of probability, *Transactions of the American Mathematical Society*, **63** (1948), pp. 422—438.
[3] W, FELLER, On the integral equation of renewal theory, *Annals of Mathematical Statistics*, **12** (1941), pp. 243—267.
[4] W. FELLER, *On probability problems in the theory of counters*, Courant Anniversary Volume (New-York, 1948), pp. 105—115.
[5] L. KOSTEN, On the frequency distribution of the number of discharges counted by a Geiger—Müller counter in a constant interval, *Physica*, **10** (1943), pp. 749—756.
[6] C. LEVERT and W. L. SCHEEN, Probability fluctuation of discharges in a Geiger—Müller counter produced by cosmic radiation, *Physica*, **10** (1943), pp. 225—238.
[7] A. RÉNYI, On some problems concerning Poisson processes, *Publicationes Mathematicae (Debrecen)*, **2** (1951), pp. 66—79.
[8] L. TAKÁCS, Occurrence and coincidence phenomena in case of happenings with arbitrary distribution law of duration, *Acta Math. Acad. Sci. Hung.*, **2** (1951), pp. 275—298.

# О ПРОЦЕССАХ ПОРОИСШЕСТВИЯ, ПОРОЖДЕННЫХ ПРОЦЕССОМ ПУАССОНА

Л. Такач (Будапешт)

(Резюме)

Рассмотрим $m$ независимых друг от друга процессов происшествия в моментах $0 \leqq u < \infty$. Каждый из этих процессов определяется следующим образом: Все те события (наступающие в интервале времени $0 \leqq u < \infty$) процесса Пуассона с плотностью событий $p$, во время наступления которых нет в ходе происшествие, пускают в ход происшествия, сроков $\chi_1, \chi_2, \chi_3, \ldots$ соответственно. Предположим, что сроки $\chi_n$— независимые случайные величины с общей функцией распределения $H(x)$.

Говорим, что система состоящая из $m$ процессов происшествия находится в момент $u$ в состоянии $E_j$ $(j = 0, 1, 2, \ldots, m)$, если число текущих происшествий равно $j$.

В работе исследуется распределение числа переходов $E_0 \to E_1$ наступающих в интервале времени $t$, и распределение времени проведенного вне состоянии $E_0$. Решим поставленные проблемы впервые для интервала времени $(0, t)$, а потом для $(u, u+t)$, при предельном переходе $u \to \infty$.

В предыдущей работе мы определили средние этих распределений, а теперь, с помощью нового метода, выведем из функции средних функции распределений. Определение этих функций распределения приводится к определению распределения интервалов времени между последовательными переходами $E_0 \to E_1$. Если ета функция распределения обозначается через $G(x)$, а ее преобразованная Лапласа—Стильтьеса через $\varphi(s)$, то справедливо соотношение

$$\varphi(s) = 1 - \frac{p^m}{mp+s}\left[\int_0^\infty e^{-st}(f(t))^m\, dt\right]^{-1},$$

где $f(t)$ есть однозначно определенное решение следующего интегрального уравнения типа Волтерра:

$$f(t) = p - p\int_0^t f(t-x)\,[1-H(x)]\, dx.$$

Во второй части работы исследуется процесс Пуассона плотности вероятностей $\lambda$, рассмотренная в моментах времени $0 \leqq u < \infty$. Предполагается, что все события пускают в ход происшествия, сроков $\chi_1, \chi_2, \chi_3, \ldots$ соответственно, где $\chi_n$ последовательность независимых, одинаково распределенных случайных величин; обозначим через $H(x)$ их общую функцию распределения. Говорим, что система находится в состоянии $E_j$ $(j = 0, 1, 2, \ldots)$, если число текущих происшествий равно $j$. Этот процесс получается из вышеописанного, если провести предельный переход $m \to \infty$ и $p \to 0$ при $mp = \lambda$. Исследуем также в этом случае распределение переходов $E_0 \to E_1$ и распределение времени проведенного вне состоянии $E_0$. Определение этих распределений приводится к определению распределения интервалов времени между последовательными переходами $E_0 \to E_1$. Обозначив эту функцию распределения через $G(x)$, получим для ее преобразованной Лапласа—Стильтьеса

$$\varphi(s) = 1 - \frac{1}{\lambda+s}\left[\int_0^\infty e^{-st-\lambda\int_0^t [1-H(x)]\,dx}\, dt\right]^{-1}.$$

Вышеуказанные процессы играют важную роль при решении проблем, связанных с теорией счетчиков частиц.

# INVESTIGATION OF WAITING TIME PROBLEMS BY REDUCTION TO MARKOV PROCESSES

By

LAJOS TAKÁCS (Budapest)

(*Presented by* A. RÉNYI)

## Introduction

By *waiting time* problems we mean questions connected with the following problem:

Let us consider a service (a counter with a single server) at which customers are arriving in the moments $0 \leqq t_1 < t_2 < \cdots < t_n < \cdots$. The sequence $\{t_n\}$ of the moments of arrivals is determined by some probability law. The arriving customers will be attended by the server in the order of their arrivals as follows: if a customer arrives at the service in a moment when nobody is waiting or attended, then he will be attended immediately. If the server is busy when a customer arrives, then this person starts to be served at the moment when the preceding finished. We suppose that the durations of the service time of the customers are random variables. We denote them by $\chi_1, \chi_2, \ldots, \chi_n, \ldots$.

If the probability laws governing the sequences of random variables, $\{t_n\}$ and $\{\chi_n\}$, furthermore, the state of the system in the moment $t=0$ are known (the service is free or in what time will it be free), then the mentioned stochastic process is uniquely defined.

We may investigate the following questions connected with the mentioned stochastic process: What is the waiting time of the $n$-th arriving customer? What is the waiting time of a customer eventually arriving at the moment $t$? To what limit does the distribution function of the waiting time converge, when $n \to \infty$ and $t \to \infty$, respectively? What is the number of waiting customers in a given moment? What kind of law is valid for the service periodes and breaks?

We shall deal with these questions in the case when $t_1, t_2, \ldots, t_n, \ldots$ are the moments of events occurring in a Poisson process. We generally suppose the Poisson process homogeneous in time, i. e. the density of events

$\lambda(t) \equiv \lambda$ (constant), but we give the results also for the inhomogeneous case, when we can do it. Concerning the sequence $\{\chi_n\}$ of service times we shall suppose that they are equidistributed mutually independent positive random variables.

There is also another interesting case. Suppose the differences $t_n - t_{n-1}$ $(n = 1, 2, \ldots;\ t_0 = 0)$ (the time intervals between two consecutive arrivals) are mutually independent positive random variables with the same (arbitrary) distribution law. We shall not consider this case in details. Originally we have also dealt with this question in our work, however, the paper of D. V. LINDLEY which appeared in the meantime made our investigations superfluous.

In the case when the sequence $\{t_n\}$ consists of the moments of the events of a Poisson process with density $\lambda$, A. JA. KHINTCHINE has determined the characteristic function of the waiting time, for the stationary state. This result of KHINTCHINE may be obtained simply by our methods. Further we shall discuss the problem also for non-stationary state, which seemed hitherto to be a very complicated case. Namely, the process is in general a non-Markov process and the "markovization" seems to lead to the solution of infinite system of integro-differential equations. However, we have stated that the process is also of Markov type, if we characterize the state by a real number: the waiting time. In this case the problem can be described by a single integro-differential equation.

We have to mention here the case of a many servers system too. In this case the waiting time problem has been dealt with only under simplifying assumptions, essentially reducing the problem to the case of a single server. Here A. K. ERLANG, F. POLLACZEK and others obtained several results.

## § 1. Formulation of the problem

Let us consider a stochastic process of the Poisson type in the moments $t \geq 0$. Let $\lambda(t)$ be the density of the occurrence of the events, where $\lambda(t)$ is a real-valued non-negative, continuous and bounded function of the time parameter $t$. In the case of a Poisson process we suppose that *the numbers of the events occurring in non-overlapping time intervals are mutually independent random variables, moreover, that the probability that in the time interval* $(t, t+\Delta t)$ *an event occurs, is* $\lambda(t)\Delta t + o(\Delta t)$ *and that of occurrence of more than one event is* $o(\Delta t)$.

Let us suppose that in the moments when an event occurs in the Poisson process, a customer arrives at a service. If there is no person waiting, then he will be served immediately. If there are waiting customers, then he must wait until the persons arrived earlier will have been attended and he begins being served in the moment when the preceding person finished. Let the durations $\chi_1, \chi_2, \ldots, \chi_n, \ldots$ of the service times be mutally independent

random variables with the same distribution function $P(\chi_n \leqq x) = H(x)$. We suppose $\chi_n$ to be positive, i. e. $H(0) = 0$.

This process occurs in several cases in practice, e. g. in the case of callings arriving at a telephone centre, landing of aeroplanes, waiting of persons before a cash-box and other waiting processes.

The above terminology has been chosen for the sake of clarity. Evidently, the problem may be also formulated abstractly.

Let the duration of waiting of a person arriving eventually in the moment $t$ be denote by the random variable $\eta(t)$. The process described by the probability function $\eta(t)$ will be called *homogeneous in time,* if $\lambda(t) \equiv \lambda$ (a constant not depending on $t$). If $\lambda(t)$ depends effectively on the time parameter $t$, then we call the process *inhomogeneous in time.*

Let $P(\eta(t) \leqq x) = F(t, x)$, i. e. $F(t, x)$ is the probability that the waiting time of a person arriving in the moment $t$ will not exceed $x$. Further let $P(\eta(t_n - 0) \leqq x) = F_n(x)$, i. e. $F_n(x)$ is the probability that the waiting time of the $n$-th arriving person is at most $x$. Here $t_n$ $(n = 1, 2, \ldots)$ denotes the moment of arriving of the $n$-th person.

First we shall deal with the determination of $F(t, x)$ and $F_n(x)$. Then we investigate the condition under which the limiting distributions $\lim\limits_{t\to\infty} F(t, x)$ and $\lim\limits_{n\to\infty} F_n(x)$ exist and show how they can be obtained. It will be shown that, under general conditions, these limiting distributions exist and they agree with each other. This common limiting distribution function gives the stationary solution, i. e. the probability that — when the process is going on since an infinitely long time — the waiting time in any moment or the waiting time of any person does not exceed $x$.

Having discussed these problems, we shall deal with the determination of the duration of service periodes and of the number of waiting persons.

The theory of waitings was discussed by several authors. We mention the book by Th. C. FRY [8] in which, using the results of A. K. ERLANG, and supposing a stationary state he has dealt with the laws for waitings. Namely, he has determined the average waiting time in case of service times with constant and exponential distribution, and the probability that the number of the waiting persons in a given moment equals $j$ $(j = 0, 1, 2, \ldots)$. (In the case of exponentially distributed service times FRY [8] considers also the more general problem of a many servers system. Here he gives also the distribution function of the waiting time.) In his paper [14] F. POLLACZEK discusses the case when there are $\nu \geqq 1$ servers and the servers attend only every $\nu$-th arriving person. He deduces the distribution function of waiting times when there is a given number of persons arriving in a given time interval at the service and the service times are distributed arbitrarily. In this paper POLLACZEK gives several asymptotic formulae for the distribution func-

tion and their expected values. From among the other works of POLLACZEK we mention only [15] in which he considers some waiting problems arising in connection with the landing of aeroplanes. In his work [11] A. N. KOLMOGOROV discusses the case when the moments of the arriving and leaving persons form a Poisson process, i. e. the duration of service times is exponentially distributed. Far-reaching results are due to A. J. KHINTCHINE [9], dealing with the stated problem in the stationary case with arbitrary service times: namely the determination of the characteristic function of waiting time. Further we mention the work of D. V. LINDLEY [13] containing KHINTCHINE's result on the stationary state as a special case. Furthermore, the work of D. G. KENDALL [10] is to be mentioned in which a very detailed bibliography of the theory of waiting time problems may be found.

In the sequel we shall discuss a problem which is more general than the preceding ones. Namely, we shall treat the non-stationary process, when it is homogeneous in time, and prove results for the inhomogeneous process too. We shall give conditions for the existence of a stationary state and the equilibrium distributions.

We think we have solved the general problem more simply and shortly than many authors the special cases of the problem. This simple solution may be explained as follows: Up to now the general opinion was (see e. g. W. FELLER [5], p. 246) that the waiting time problem leads to a Markov process only in the special case when the service time is distributed exponentially. In this case let the state $E_j$ denote that the number of persons waiting and being served is equal to $j$. Then the changes of states actually form a Markov process. If, however, the service time has not an exponential distribution, then this process ceases to be of Markov type. It becomes of Markov type by extending the notion of the "state" by including in the definition of the state a continuous variable, namely the time passed since the begin of service of the person just being served. In this case, however, quoting FELLER, "The new equations can usually be written down but they are so complicated that it is doubtful whether anything is gained". However, we have noticed that the process may be considered as of Markov type even if we characterize the state of the system by a single variable, the waiting time of the person arriving eventuelly in the moment considered. This description may be carried out equally in the cases of service times with exponential and arbitrary distribution, respectively. It is a great advantage that in this procedure the states $E_j$ do not play any role. Consequently, we have to deal only with a single integro-differential equation instead of a system with infinitely many equations.

## § 2. The distribution function of the waiting time

Let $\eta(t)$ be the waiting time of a person arriving eventually at the service in the moment $t$. If in the moment $t=0$ the server is free, then $\eta(0)=0$. If in the moment $t=0$ the server is not free, then we denote by $\eta(0)=\eta_0$ the random variable giving the moment when the server ceases to be busy. If in a given moment $\eta(t)=0$, then it remains zero until a person arrives at the service.

If $\{t_n\}$ and $\{\chi_n\}$ denote the sequence of moments of arrivals and the sequence of the corresponding service times, respectively, then the value of $\eta(t)$ will have the jump $\chi_n$ $(n=1,2,\ldots)$ in the moments $t_n$ $(n=1,2,\ldots)$. In the meantime the value of $\eta(t)$ decreases linearly, but, if it assumes the value 0, then it remains zero until the moment of arrival of the next person.

Thus $\eta(t)$ is defined as follows: $\eta(0)=\eta_0$ and if $t_n\leq t<t_{n+1}$ $(n=0,1,2,\ldots;$ $t_0=0)$, then

$$\eta(t)=\begin{cases}\eta(t_n)-(t-t_n) & \text{if} \quad \eta(t_n)>t-t_n,\\ 0 & \text{if} \quad \eta(t_n)\leqq t-t_n\end{cases} \tag{1}$$

and if $t=t_n$, then

$$\eta(t_n)=\eta(t_n-0)+\chi_n. \tag{2}$$

If we know the probability laws governing the sequences $\{t_n\}$ and $\{\chi_n\}$ of random variables, then by the initial condition $\eta(0)=\eta_0$ and by equations (1) and (2) $\eta(t)$ may be determined in any moment $t$.

We emphasize that the probability function $\eta(t)$ is defined for all $t\geqq 0$ and $\eta(t_n-0)$ yields the waiting time of the $n$-th arriving person. The course of $\eta(t)$ is illustrated by Fig. 1.

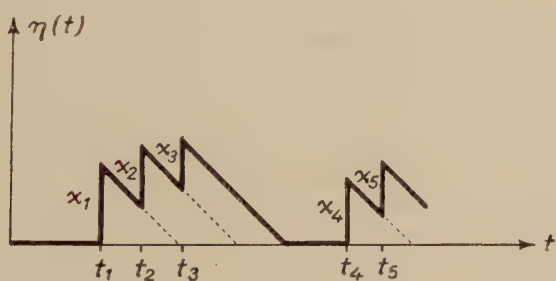



Fig. 1

If $t_n$ agree with the moments of the events in a Poisson process and $\chi_n$ are mutually independent random variables, then the probability function $\eta(t)$ describes a Markov process. In fact, if we know the value of $\eta(t)$ in a given moment, this determines the future stochastic behaviour of $\eta(t)$ completely.

If the moments $t_n$ form no Poisson process, but the time differences $t_n-t_{n-1}$ $(n=1,2,\ldots)$ are mutually independent random variables of the same

distribution, then the process $\eta(t)$ will not be of Markov type, but the moments $t_n$ will be the Markov points (or, in an other nomenclature, "the regeneration points") of the process. Thus, also in this case we can reduce the process to the study of Markov processes.

In the sequel we shall suppose that $t_n$ denote the moments of the events in a Poisson process of density $\lambda(t)$ and $\chi_n$ are equidistributed mutually independent random variables with the common distribution function $H(x)$. In general, we shall suppose that $\eta(0)=\eta_0$ is arbitrary with the distribution function $P(\eta_0 \leqq x)=F_0(x)$ or, in particular, $\eta_0 \equiv 0$ with $F_0(x)=0$ for $x<0$ and $F_0(x)=1$ for $x \geqq 0$. Now, for the distribution function $P(\eta(t) \leqq x)=F(t,x)$ we prove the folloving

THEOREM 1. *The distribution function $F(t,x)$ of the random variable $\eta(t)$ satisfies the integro-differential equation*

$$\frac{\partial F(t,x)}{\partial t}=\frac{\partial F(t,x)}{\partial x}-\lambda(t)F(t,x)+\lambda(t)\int_0^x H(x-y)\,d_yF(t,y); \tag{3}$$

*all the derivatives occurring in* (3) *are either right-hand derivatives* ($x \geqq 0$) *or left-hand derivatives* ($x>0$). *The distribution function $F(t,x)$ has a jump of value $F(t,0)$ at $x=0$ and for $x>0$, is continuous for all $t$. Under the initial condition $F(0,x)=1$* ($x \geqq 0$) (3) *has a unique solution $F(t,x)$.*

PROOF. For $0<\Delta t$ we have the equation

$$F(t+\Delta t,x)=(1-\lambda(t)\Delta t)\,F(t,x+\Delta t)+\lambda(t)\Delta t\int_0^x H(x-y)\,d_yF(t,y)+o(\Delta t). \tag{4}$$

Indeed, the event $\eta(t+\Delta t) \leqq x$ may happen in several mutually excluding ways, namely:

1. In the interval $(t, t+\Delta t)$ no event occurs (the probability is $1-\lambda(t)\Delta t+o(\Delta t)$), then we have $\eta(t) \leqq x+\Delta t$ with the probability $F(t,x+\Delta t)$.

2. In the interval $(t, t+\Delta t)$ one event occurs; its probability is $\lambda(t)\Delta t+o(\Delta t)$ and if $\eta(t)=y\,(0 \leqq y<x)$, then we must have $\chi \leqq x-y$ whose probability is $H(x-y)$ and the distribution function of $y$ is $F(t,y)$.

3. In the interval $(t, t+\Delta t)$ more than one event occurs the probability whereof is $o(\Delta t)$.

Later we shall show that for $x>0$ the distribution function $F(t,x)$ is a continuous function of $x$. At $x=0$, $F(t,x)$ has in general a jump of value $F(t,0)$. So we have for $\Delta t>0$

$$F(t,x+\Delta t)=F(t,x)+\frac{\partial F(t,x)}{\partial x}\Delta t+o(\Delta t) \tag{5}$$

where $\partial F/\partial x$ denotes a right-hand derivative which exists for $x \geqq 0$. Putting

(5) in (4), the limit-transition $\Delta t \to 0$ yields for $x \geqq 0$

$$\frac{\partial F(t,x)}{\partial t} = \frac{\partial (F,t,x)}{\partial x} - \lambda(t) F(t,x) + \lambda(t) \int_0^x H(x-y)\, d_y F(t,y)$$

which agrees with (3). This equation shows that $F(t,x)$ is differentiable with respect to $t$ from the right and the derivative agrees with the right-hand side of the equation. Similarly, it may be shown that for $x > 0$ the left-hand derivative of $F(t,x)$ with respect to $t$ also exists and it agrees with the right-hand side of (3), the only difference is that here $\partial F/\partial x$ denotes the left-hand derivative with respect to $x$. For such values $x$ for which the right and left-hand derivatives of $F(t,x)$ with respect to $x$ coincide (i. e. for almost every $x$), $F(t,x)$ is differentiable with respect to $t$ and its derivative agrees with the right-hand side of (3).

The process described by the probability function $\eta(t)$ is a so-called Markov process of mixed type, since it consists of a mixture of discontinuous and continuous state-changings. Processes of this type have been dealt with by W. FELLER [4], he has namely shown that, under certain conditions, a unique solution $F(t,x)$ exists which is a distribution function in $x$. FELLER also gives the solution in the form of uniformly convergent infinite series. Unfortunately, his general discussion can not be immediately applied to this special case.

Instead of the solution $F(t,x)$ we pass now to the study of its Laplace—Stieltjes transform. In this way we shall not only prove our Theorem but also obtain the solution itself. Let us introduce the following functions converging for $\Re(s) \geqq 0$:

(6) $$\Phi(t,s) = \int_0^\infty e^{-sx}\, dF(t,x)$$

and

(7) $$\psi(s) = \int_0^\infty e^{-sx}\, dH(x).$$

Let us form the Laplace—Stieltjes transform of equations (3) on both sides. Then we obtain for $\Re(s) \geqq 0$

(8) $$\frac{\partial \Phi(t,s)}{\partial t} = \Phi(t,s)[s - \lambda(t) + \lambda(t)\psi(s)] - sF(t,0).$$

It is easy to prove that the derivative $\partial \Phi/\partial t$ exists for all values of $t$ and $F(t,0) = \lim\limits_{x \to +0} F(t,x)$, the jump of $F(t,x)$ at $x = 0$. The coefficients of this differential equation are continuous functions of $t$ and $\Phi(0,s) = 1$. Under this initial conditions, (8) has, for $\Re(s) \geqq 0$ as well known, a unique solution

(9) $$\Phi(t,s) = e^{st - [1-\psi(s)]\Lambda(t)} \left[ 1 - s \int_0^t e^{-su + [1-\psi(s)]\Lambda(u)} F(u,0)\, du \right]$$

where

$$\Lambda(t)=\int_0^t \lambda(u)\,du. \tag{10}$$

Knowing $\Phi(t,s)$, the distribution function $F(t,x)$ may be determined uniquely, as it is known by the theory of the Laplace transform. So formula (9) determines $F(t,x)$ uniquely, provided that $F(u,0)$ may be given uniquely for the values $0\leqq u\leqq t$. The fact that $F(t,0)$ can be determined uniquely for every $t$ may be elucidated by regarding the process from the point of view of the server. The time of the server consists of alternating breaks and service periodes and $F(t,0)$ is the probability that there is a break in the moment $t$. Knowing the distribution function of the duration of service periodes, $F(t,0)$ may be determined uniquely. In case of a process inhomogeneous in time, this calculation is very complicated but it may be performed easily for homogeneous processes, which also will be done in the next chapter.

It remains to prove that the distribution function $F(t,x)$ is continuous with respect to $x$ for every value $t$. $F(t,x)$ may be presented also in the following form:

$$\begin{aligned} F(t,x)=e^{\Lambda(t)}\Bigg[1+\int_0^t F(t-u,0)\,\lambda(t-u)\,e^{\Lambda(t-u)}H(u+x)\,du+ \\ +\int_0^t\Bigg(\int_0^{x+u} F(t-u,x+u-y)\,\lambda(t-u)\,e^{\Lambda(t-u)}\,dH(y)\Bigg)du\Bigg]. \end{aligned} \tag{11}$$

In fact, $\eta(t)\leqq x$ if 1) in the interval $(0,t)$ no event occurs, 2) in the moment $t-u$ (where $0\leqq u\leqq t$) a service begins without waiting with a duration $\leqq u+x$ and in the interval $(t-u,t)$ no event occurs, 3) in the moment $t-u$ (where $0\leqq u\leqq t$) a person arrives at the service and his waiting time is smaller than $x+u-y$, where $y$ denotes the service time and in the interval $(t-u,t)$ nobody arrives at the service.

From the representation (11), which is at the same time a recurrent formula for determining $F(t,x)$, we can easily show the continuity of $F(t,x)$ for $x>0$.

By our definition $F(t,x)$ is a non-decreasing function of $x$. $F(t,\infty)=1$ follows from the fact that for the Laplace—Stieltjes transform of $F(t,x)$ we have $\Phi(t,0)=1$.

Owing to what has been said above we can also state the following

THEOREM 2. *The Laplace—Stieltjes transform*

$$\Phi(t,s)=\int_0^\infty e^{-sx}\,dF(t,x) \tag{12}$$

*of the solution* $F(t,x)$ *of the integro-differential equation* (1) *in Theorem* 1

*may be presented in the following form:*

$$\Phi(t, s) = e^{st-[1-\psi(s)]\Lambda(t)}\left[1 - s\int_0^t e^{-su+[1-\psi(s)]\Lambda(u)} F(u, 0)\, du\right] \tag{13}$$

*where $F(u, 0)$ denotes the probability that in the moment $u$ the service is free.*

We should have dealt similarly with the case when $\eta_0$ is not identically zero but it is an arbitrary random variable. If we suppose that the distribution function $F_0(x)$ of $\eta_0$ is continuous for $x > 0$, then the above discussion can be applied step by step with the sole modification that in (13) the initial condition is $\Phi(0, s) = \Phi_0(s)$ instead of $\Phi(0, s) = 1$, where $\Phi_0(s)$ denotes the Laplace—Stieltjes transform of $F_0(x)$ and $F(t, 0)$ is also to be substituted by an expression corresponding to the new case.

Next we pass to the determination of the limiting distribution $\lim_{t\to\infty} F(t, x) = F^*(x)$. Here we shall prove the following theorems:

THEOREM 3. *Let the expected value of the random variable $\chi_n$ be*

$$\alpha = \int_0^\infty x\, dH(x) = \int_0^\infty [1 - H(x)]\, dx \tag{14}$$

*and let*

$$\lim_{t\to\infty} \lambda(t) = \lambda \tag{15}$$

*be a positive constant. Now if $\lambda\alpha < 1$, then the limiting distribution function $\lim_{t\to\infty} F(t, x) = F^*(x)$ exists, is independent of the initial distribution $F_0(x)$ and is uniquely determined by the equations*

$$F^*(0) = 1 - \lambda\alpha \tag{16}$$

*and*

$$\frac{dF^*(x)}{dx} = \lambda\left[F^*(x) - \int_0^x H(x-y)\, dF^*(y)\right] \tag{17}$$

*where $\frac{dF^*(x)}{dx}$ denotes the right-hand derivative.*

*If $\lambda\alpha \geqq 1$, then the limiting distributon $F^*(x)$ does not exist, howewer, $\lim_{t\to\infty} F(t, x) = 0$ for every $x$.*

THEOREM 4. *If $\lim_{t\to\infty} \lambda(t) = \lambda$ and $\lambda\alpha < 1$, then the limiting distribution $\lim_{t\to\infty} F(t, x) = F^*(x)$ exists, is independent of the initial distribution $F_0(x)$ and is uniquely determined by the Laplace—Stieltjes transform*

$$\Phi^*(s) = \int_0^\infty e^{-sx}\, dF^*(x), \quad (\Re(s) \geqq 0), \tag{18}$$

*satisfying*

$$\Phi^*(s) = \frac{1-\lambda\alpha}{1-\lambda\dfrac{1-\psi(s)}{s}}. \tag{19}$$

PROOF. First we shall establish that, independently of the initial distribution $F_0(x)$, the limit $\lim_{t\to\infty} F(t,0) = 1-\lambda\alpha$ exists if $\lambda\alpha < 1$ and $\lim_{t\to\infty} F(t,0) = 0$ if $\lambda\alpha \geqq 1$. If $\eta_0 \equiv 0$ and $\lambda(t) \equiv \lambda$ (constant), then we shall show in the next chapter that the limit $\lim_{t\to\infty} F(t,0) = F^*(0) = 1-\lambda\alpha$ is valid if $\lambda\alpha < 1$. If $\eta_0$ is arbitrary, then we consider the moment when the initial reservation of the service ceases. Then, no matter how many persons are waiting there, the probability of the state $E_0$ will be the same when $t\to\infty$, independently of the initial state, i. e. also for an arbitrary $\eta_0$ the limit $\lim_{t\to\infty} F(t,0) = F^*(0) = 1-\lambda\alpha$ is valid. Namely, in the next chapter we shall show that the state $E_0$ is ergodic, hence, our statement will be a consequence of the theory of Markov chains. If $\lim_{t\to\infty} \lambda(t) = \lambda$, then, from some sufficiently large value of $t$ on, the relation $\lambda_1 \leqq \lambda(t) \leqq \lambda_2$ holds for arbitrarily given values of $\lambda_1$ and $\lambda_2$ around $\lambda$, and independently of the initial distribution, $F_{\lambda_2}(t,0) \leqq F(t,0) \leqq F_{\lambda_1}(t,0)$ is valid for sufficiently large values of $t$, where $F_\lambda(t,0)$ denotes the probability of the state $\eta(t)=0$ in case of a process of density $\lambda$. This follows from the ergodicitiy of the state $E_0$, using the same argument as above. Consequently, if $\lambda_2\alpha < 1$, then $1-\lambda_2\alpha \leqq \lim_{t\to\infty} F(t,0) \leqq 1-\lambda_1\alpha$. Since $\lambda_1$ and $\lambda_2$ may be chosen arbitrarily near to $\lambda$, therefore by $\lambda\alpha < 1$ we have $\lim_{t\to\infty} F(t,0) = F^*(0) = 1-\lambda\alpha$. If $\lambda\alpha \geqq 1$, then, as it will be seen, in case of $\lambda(t) \equiv \lambda$ (constant) we have $\lim_{t\to\infty} F(t,0) = 0$ which is true in general too.

In the sequel, first we shall suppose $\eta_0 \equiv 0$, i. e. $F_0(x) \equiv 1$ if $x \geqq 0$.

Now let $s = -i\omega$ in formula (13) (where $\omega$ is real). So we obtain the characteristic function of the distribution function $F(t,x)$ and denote it by $\Phi_t(\omega)$. Then

$$\Phi_t(\omega) = e^{-i\omega t-[1-\psi(-i\omega)]\Lambda(t)}\left[1+i\omega\int_0^t e^{i\omega u+[1-\psi(-i\omega)]\Lambda(u)}F(u,0)\,du\right]. \tag{20}$$

For every $t$, $\Phi_t(\omega)$ is the characteristic function of a non-negative random variable. So we may apply a theorem due to A. ZYGMUND [18] stating that if, in an arbitrary interval containing the point $\omega=0$, we have $\lim_{t\to\infty}\Phi_t(\omega) = \Phi(\omega)$ and $\Phi(\omega)$ is continuous at $\omega=0$, then the limiting distribution function $\lim_{t\to\infty} F(t,x) = F^*(x)$ exists and its characteristic function is equal to $\Phi(\omega)$.

Now it will be shown that for the values $\omega$ satisfying $\Re[\psi(-i\omega)] < 1$ we have

$$\lim_{t\to\infty} \Phi_t(\omega) = \frac{F^*(0)}{1+\lambda\dfrac{1-\psi(-i\omega)}{i\omega}}. \tag{21}$$

Let $\omega$ be a fixed value for which $\Re[\psi(-i\omega)] < 1$ and, for the sake of brevity, let $\varphi(t) = i\omega t + [1-\psi(-i\omega)]\,\Lambda(t) = i\omega t + \gamma\Lambda(t)$ where $\Re(\gamma) > 0$. Then

$$\Phi_t(\omega) = e^{-\varphi(t)}\left[1 + i\omega\int_0^t e^{\varphi(u)} F(u,0)\,du\right].$$

Now $\lim\limits_{t\to\infty} e^{-\varphi(t)} = 0$ and we have to prove

$$\lim_{t\to\infty} e^{-\varphi(t)}\int_0^t e^{\varphi(u)} F(u,0)\,du = \frac{F^*(0)}{\varphi^*} \tag{22}$$

where $\varphi^* = \lim\limits_{t\to\infty} \varphi'(t) = i\omega + [1-\psi(-i\omega)]\lambda$. For a fixed value of $T$ we have

$$\lim_{t\to\infty} e^{-\varphi(t)}\int_0^T e^{\varphi(t)} F(u,0)\,du = 0, \tag{23}$$

so we have to verify only that, for a sufficiently large $t$,

$$\left| e^{-\varphi(t)}\int_T^t e^{\varphi(u)} F(u,0)\,du - \frac{F^*(0)}{\varphi^*}\right| \tag{24}$$

may be made arbitrarily small.

First of all we prove that, for a sufficiently large $T$,

$$|e^{-\varphi(t)}|\int_T^t |e^{\varphi(u)}|\,du < \frac{2}{\lambda\delta} \tag{25}$$

where $\delta = \Re(\gamma)$, i. e. (25) is a bounded function. In fact, we have $\Re(\varphi(t)) = \delta\Lambda(t)$ and hence

$$|e^{-\varphi(t)}|\int_T^t |e^{\varphi(u)}|\,du = e^{-\delta\Lambda(t)}\int_T^t e^{\delta\Lambda(u)}\,du \leqq \frac{2e^{-\delta\Lambda(t)}}{\lambda}\int_T^t e^{\delta\Lambda(u)}\lambda(u)\,du \leqq \frac{2}{\lambda\delta} \tag{26}$$

where $T$ has been chosen such that for $t \geqq T$ we have $\lambda(t) \geqq \lambda/2$.

Now the inequality

$$\left| e^{-\varphi(t)}\int_T^t e^{\varphi(u)} F(u,0)\,du - \frac{F^*(0)}{\varphi^*}\right| \leqq \left| e^{-\varphi(t)}\int_T^t e^{\varphi(u)}[F(u,0)-F^*(0)]\,du\right| + \tag{27}$$

$$+ F^*(0)\left| e^{-\varphi(t)}\int_T^t e^{\varphi(u)}\,du - \frac{1}{\varphi^*}\right|$$

is valid. The first term of the right-hand side is $\leq \frac{2}{\lambda\delta} \max\limits_{(T,t)} |F(u,0)-F^*(0)|$ which may be made arbitrarily small for sufficiently large values $T$. Making use of the equations

$$e^{-\varphi(t)}\int\limits_T^t e^{\varphi(u)}\varphi'(u)\,du = 1-e^{\varphi(T)-\varphi(t)} \tag{28}$$

and $\varphi'(u)=\varphi^*+\gamma[\lambda(u)-\lambda]$, the second term may be written as

$$e^{-\varphi(t)}\int\limits_T^t e^{\varphi(u)}\,du-\frac{1}{\varphi^*}=-\frac{e^{\varphi(T)-\varphi(t)}}{\varphi^*}-\frac{\gamma e^{-\varphi(t)}}{\varphi^*}\int\limits_T^t e^{\varphi(u)}[\lambda(u)-\lambda]\,du. \tag{29}$$

For sufficiently large $t$ and $T$, both the first and the second term on the right-hand side may be made arbitrarily small. For $\lambda(u)-\lambda$ may be made arbitrarily small for $u\geqq T$ and (25) is valid. Thus (21) has been established.

It remains to show that there is a neighbourhood of $\omega=0$ such that $\Re\left(\int\limits_0^\infty e^{i\omega x}\,dH(x)\right)<1$. If $H(x)$ is a degenerated distribution, i. e. if $\chi_n$ is equal to a constant $\alpha$ different from zero, then this is obvious. If, on the other hand, $H(x)$ is no degenerated distribution, then, from some result discussed in the book of A. N. KOLMOGOROV and B. V. GNEDENKO [12, p. 60], it follows that there is a number $\omega_0\neq 0$ such that in the interval $(-\omega_0,\omega_0)$ we have $|\psi(-i\omega)|<1$, i. e. $\Re\,\psi(-i\omega)<1$ also holds. Furthermore, we have to show that (21) is continuous in the point $\omega=0$. Since the characteristic function $\psi(-i\omega)$ is uniformly continuous for all $\omega$ and the limit

$$\lim_{\omega\to 0}\frac{1-\psi(-i\omega)}{i\omega}=-\alpha \tag{30}$$

exists, (21) is also continuous at $\omega=0$. Thus we have shown that the limiting distribution $F^*(x)$ exists. But if the limiting distribution $F^*(x)$ exists, then, from (3) it follows that $F^*(x)$ satisfies the integro-differential equation (17).

Evidently, the Laplace—Stieltjes transform of $F^*(x)$ is

$$\Phi^*(s)=\int\limits_0^\infty e^{-sx}\,dF^*(x)=\frac{F^*(0)}{1-\lambda\dfrac{1-\psi(s)}{s}}=\frac{1-\lambda\alpha}{1-\lambda\dfrac{1-\psi(s)}{s}}. \tag{31}$$

Here $\Phi^*(0)=1$, since

$$\lim_{s\to 0}\frac{1-\psi(s)}{s}=\int\limits_0^\infty x\,dH(x)=\alpha. \tag{32}$$

If $\lambda\alpha\geqq 1$, then $\lim\limits_{t\to\infty}F(t,0)=0$ and it is easy to prove that $\lim\limits_{t\to\infty}\Phi_t(\omega)=0$ whence $\lim\limits_{t\to\infty}F(t,x)=0$ for all $x$.

If the distribution function $F_0(x)$ of $\eta_0$ is arbitrary, then the limiting distribution $\lim_{t\to\infty} F(t, x)$ also exists and agrees with $F^*(x)$. Indeed, the state $E_0$ being ergodic, the system reaches the state $E_0$ with in a finite time with probability one, and, as proved above, starting from the state $E_0$, $F(t, x)$ tends to the limiting distribution $F^*(x)$ when $t\to\infty$.

Formula (31) is equivalent to the solution given for the stationary case by A. JA. KHINTCHINE [9].

REMARK. If, following KHINTCHINE, we suppose that, in case $\lambda\alpha<1$, the stationary state with a distribution function $F^*(x)$ exists, then we may also obtain equation (31) starting from equation (3). For, from (3) the validity of (17) follows immediately. Using it, for the Laplace—Stieltjes transform $\Phi^*(s)$ of $F^*(x)$ we obtain

$$\Phi^*(s)=\frac{F^*(0)}{1-\lambda\dfrac{1-\psi(s)}{s}}.$$

Since $F^*(x)$ is a distribution function, we must have $\lim_{s\to 0}\Phi^*(s)=1$ and, on the other hand, $\lim_{s\to 0}[1-\psi(s)]/s=\alpha$ holds, so we get

$$F^*(0)=1-\lambda\alpha.$$

Now let us suppose that the moments $t_n$ agree with the moments of the events in a homogeneous Poisson process of density $\lambda$. Let $F_n(x)$ denote the distribution function of the waiting time of the $n$-th person, i. e. let $P(\eta(t_n-0)\leqq x)=F_n(x)$. If the original occupation of the service is $\eta_0$ with the distribution function $P(\eta_0\leqq x)=F_0(x)$, then the terms of the sequence $\{F_n(x)\}$ may be obtained successively by the following recurrence formula:

$$F_n(x)=\int_0^\infty K(x, y)\,dF_{n-1}(y) \qquad (n=1,2,\ldots) \tag{33}$$

where

$$K(x, y)=H(x-y)+\int_{x-y}^\infty e^{-\lambda(u+y-x)}\,dH(u). \tag{34}$$

This follows simply from the fact that the sequence $\{\eta(t_n-0)\}$ of random variables forms a Markov chain with the transition probabilities

$$P(\eta(t_n-0)\leqq x\,|\,\eta(t_{n-1}-0)=y)=K(x, y).$$

We investigate now the behaviour of $F_n(x)$ when $n\to\infty$. We shall prove the following

THEOREM 5. *If for the expected value $M(\chi_n)=\alpha$ we have $\lambda\alpha<1$, then, independently of the initial distribution function $F_0(x)$, the distribution functions $F_n(x)$ converge to $F^*(x)$ determined in Theorems 3 and 4 when $n\to\infty$. If $\lambda\alpha\geqq 1$, then $\lim_{n\to\infty} F_n(x)=0$ for all $x$.*

PROOF. First let us consider the case when $\eta_0 = 0$. Then it is clear that $F_{n+1}(x) \leqq F_n(x)$. In fact, $F_n(x)$ is the probability that the waiting time of the $n$-th person does not exceed $x$, if the service of the first person has begun without waiting, and $F_{n+1}(x)$ may be regarded as the probability that the waiting time of the $n$-th person does not exceed $x$ in a waiting where the service of the first person begins eventually after some waiting. Since the sequence $\{F_n(x)\}$ decreases monotonically and is bounded from below, there exists a limit $\lim\limits_{x \to \infty} F_n(x)$ for all $x$.

Now, let $G(t, x)$ denote the distribution function of the waiting time of the person arriving last in the time interval $(0, t)$. Further, let $K_t(x, y)$ be the conditional probability that a person, eventually arriving at the service in the moment $t$, has a waiting time $\leqq x$, provided that the person arriving last in the time interval $(0, t)$ had a waiting time $y$. Then we have

$$F(t, x) = \int_0^t K_t(x, y) d_y G(t, y). \tag{35}$$

It can be easily shown that

$$\lim_{t \to \infty} K_t(x, y) = K(x, y) \tag{36}$$

where $K(x, y)$ is defined by (33) and

$$\lim_{t \to \infty} G(t, y) = \lim_{n \to \infty} F_{n+1}(y). \tag{37}$$

Indeed, if we denote by $t_n$ the moment of the arrival of the person arriving last in the time interval $(0, t)$, then in case $t \to \infty$ the random variable $t - t_n$ is distributed exponentially with a density function $\lambda e^{-\lambda x}$ $(x \geqq 0)$ and this agrees with the distribution of the differences between consecutive arrivals. Further, if $t \to \infty$, then $n \to \infty$ with probability one. Hence by (35) we obtain

$$\lim_{t \to \infty} F(t, x) = \lim_{n \to \infty} \int_0^\infty K(x, y) d F_{n+1}(y) = \lim_{n \to \infty} F_{n+2}(x) \tag{38}$$

where the last equation follows from (33). We have shown before that the limit $\lim\limits_{n \to \infty} F_n(x)$ exists and so, by (38), it agrees with $\lim\limits_{t \to \infty} F(t, x)$, that is, $\lim\limits_{n \to \infty} F_n(x) = F^*(x)$ if $\lambda\alpha < 1$ and $\lim\limits_{n \to \infty} F_n(x) = 0$ if $\lambda\alpha \geqq 1$.

If $\eta_0$ is arbitrary, then a limiting distribution also exists and agrees with $F^*(x)$. For, as it will be proved in the next chapter, the state $E_0$ is ergodic. Thus, starting from an arbitrary initial state, we shall have, with probability one, a moment in which the system has the state $E_0$. After this moment the system behaves as if it started originally with the condition $\eta_0 = 0$.

EXAMPLES. 1. Let us have $\chi_n = \alpha$ (constant); then $\psi(s) = e^{-s\alpha}$ and by (19) in case $\lambda\alpha < 1$ we have

$$\Phi^*(s) = \frac{1-\lambda\alpha}{1-\lambda\dfrac{1-e^{-\alpha s}}{s}} = (1-\lambda\alpha)s\sum_{j=0}^{\infty}(-1)^j\frac{\lambda^j e^{-s\alpha j}}{(s-\lambda)^{j+1}} \tag{39}$$

whence

$$F^*(x) = (1-\lambda\alpha)\sum_{j=0}^{\left[\frac{x}{\alpha}\right]}(-1)^j e^{\lambda(x-\alpha j)}\frac{\lambda^j(x-j\alpha)^j}{j!}. \tag{40}$$

2. Let us suppose that $\chi_n$ has an exponential distribution, i. e. $H(x) = 1-e^{-x/\alpha}$ (if $x \geqq 0$); then $\psi(s) = 1/(1+\alpha s)$ and hence

$$\Phi^*(s) = \frac{(1-\lambda\alpha)(1+\alpha s)}{(1-\lambda\alpha)+\alpha s} \tag{41}$$

and

$$F^*(x) = 1-\lambda\alpha e^{-\frac{1-\lambda\alpha}{\alpha}x}. \tag{42}$$

## § 3. Investigations concerning homogeneous processes

In this chapter we always suppose that the moments $t_n$ are the moments of events in a homogeneous Poisson process of density $\lambda$. For the sake of simplicity let $\eta_0 = 0$, i. e. let the starting occupation of the service be zero. It is clear that the time of the server is composed of pauses and service periodes. Let $\vartheta_1, \vartheta_2, \ldots$ and $\xi_1, \xi_2, \ldots$ be the durations of consecutive pauses and service periodes, respectively. They are all mutually independent random variables. The $\vartheta_n$ are distributed exponentially with the distribution function $P(\vartheta_n \leqq x) = 1-e^{-\lambda x}$ for $x \geq 0$. The $\xi_n$ have a common distribution law $P(\xi_n \leqq x) = G(x)$ which will be determined below. We shall also determine $F(t, 0)$, i. e. the probability that in the moment $t$ the server is free. We shall prove that $\lim\limits_{t\to\infty} F(t, 0) = 1-\lambda\alpha$ when $\lambda\alpha < 1$ and $\lim\limits_{t\to\infty} F(t, 0) = 0$ when $\lambda\alpha \geqq 1$. Furthermore, in the stationary case ($t \to \infty$, $\lambda\alpha < 1$) we shall determine the probability that the system will stay in the state $E_j$ ($j = 0, 1, 2, \ldots$), i. e. that the number of persons waiting and being served be equal to $j$.

1. *Investigation of the service period. Determination of* $G(x)$. Suppose that a person without waiting arrives at the service. In the moment of the arrival he will be attended immediately. Let us suppose that the duration of this service time is $y$. The probability that in this time interval $n$ persons arrive at the counter is

$$e^{-\lambda y}\frac{(\lambda y)^n}{n!}, \tag{43}$$

i. e. it has a Poisson distribution. If $n=0$, then the service period consists of a single service and the distribution function of this is $H(x)$. If $n \geqq 1$, then the server after having attended the first person, starts to attend one of the waiting persons. The other waiting persons may be imagined standing aside. Having finished the service of this person and the persons arriving eventually in the meanwhile (the distribution function of this duration is $G(x)$) the server begins to serve one of the persons standing aside. This person and the other persons arriving in the meanwhile having been served again the service begins to attend one of the persons standig aside again. This will be continued until there are persons standing aside. From the point of wiew of the service it is perfectly indifferent whether the persons will be attend in order of their arrivals or not. This fact touches the persons only, namely, the distribution function of the waiting time is changed by this fact but $F(t, 0)$ and the expected waiting time remain unaltered. Let $G_n(x)$ denote the distribution function of the sums of $n$ mutually independent random variables whose distribution agrees with that of $\xi$, i. e. the $n$-fold convolution of $G(x)$. $G_n(x)$ may be calculated by the following recurrence formula:

$$G_n(x) = \int_0^x G_{n-1}(x-y)\, dG(y), \qquad (n=2, 3, \ldots) \tag{44}$$

where $G_1(x)=G(x)$. By what has been told we may write

$$G(x) = \int_0^x \sum_{n=0}^{\infty} e^{-\lambda y} \frac{(\lambda y)^n}{n!} G_n(x-y)\, dH(y). \tag{45}$$

In fact, the length of the service periode does not exceed $x$ if the service of the first arriving person lasts a time $y$ $(0 < y \leqq x)$; the distribution function of it is $H(y)$ and during his service time there are $n$ $(n=0, 1, 2, \ldots)$ arriving persons (the probability of it is given by (43)) and the service duration of these and those having arrived in the meanwhile does not exceed $x-y$, the probability whereof is $G_n(x-y)$ as seen. We note that $G_0(x)=1$ if $x \geqq 0$ and $G_0(x)=0$ if $x<0$.

Let the Laplace—Stieltjes transform of $G(x)$ be

$$\Gamma(s) = \int_0^\infty e^{-sx}\, dG(x) \tag{46}$$

which converges for $\Re(s) \geqq 0$.

*Passing from equation* (45) *to the Laplace transform, we obtain*

$$\Gamma(s) = \sum_{n=0}^{\infty} \frac{(\Gamma(s))^n}{n!} \int_0^\infty e^{-(s+\lambda)x} (\lambda x)^n\, dH(x) = \sum_{n=0}^{\infty} (-1)^n \frac{\lambda^n (\Gamma(s))^n \psi^{(n)}(s+\lambda)}{n!} = \tag{47}$$
$$= \psi[s+\lambda-\lambda\Gamma(s)],$$

*i. e. for* $\Re(s) \geqq 0$, $\Gamma(s)$ *the Laplace—Stieltjes transform of* $G(x)$ *satisfies the functional equation*

$$\Gamma(s) = \psi[s+\lambda-\lambda\Gamma(s)]. \tag{48}$$

A sketchy proof of this equation has been already given by D. G. KENDALL [10].

Now we prove the following theorem:

THEOREM 6. *The Laplace transform of the distribution function* $G(x)$ *of the service period* $\Gamma(s)$ *is the uniquely defined analytic solution of the functional equation*

$$\Gamma(s) = \psi[s+\lambda-\lambda\Gamma(s)], \tag{49}$$

*valid for* $\Re(s) \geqq 0$, *where* $\Gamma(s)$ *is subject to the condition* $\Gamma(\infty)=0$. $G(x)$ *can be determined uniquely by* $\Gamma(s)$.

*Denote by* $p^*$ *the smallest positive number for which*

$$\psi(\lambda(1-p^*)) = p^*, \tag{50}$$

*then* $\lim_{x\to\infty} G(x) = p^*$. *If* $\lambda\alpha \leqq 1$, *then* $p^*=1$ *and* $G(x)$ *is a proper distribution function, while if* $\lambda\alpha > 1$, *then* $p^* < 1$ *and* $G(x)$ *is an improper distribution function, namely, in such cases the service period can be infinite with probability* $(1-p^*)$.

PROOF. It is clear that $G(x)$ exists for each finite $x$ and, as we have shown, $\Gamma(s)$ satisfies the functional equation (49). Further it is clear that $\Gamma(s)$ is an analytic function for $\Re(s)>0$, and $\Gamma(\infty)=0$. Knowing $\Gamma(s)$, $G(x)$ can be determined uniquely. Therefore, it is sufficient to show that $\Gamma(s)$ is uniquely determined by $\Gamma(\infty)=0$ and by its analytic character. It is sufficient to know $\Gamma(s)$ only for non-negative real values $s$, because it can be uniquely extended for $\Re(s) \geqq 0$. The function $\psi(s)$ decreases monotonically for $0 \leqq s < \infty$; $\psi(0)=1$ and $\psi(\infty)=0$. $\psi(s)$ is differentiable for $0 \leqq s < \infty$ and $\psi'(s)$ increases monotonically; $\psi'(0) = -\alpha$ and $\psi'(\infty)=0$. Let $\Gamma(s)=x$ in (49), then $s=\Gamma^{-1}(x)$. $\psi(s)$ is a monotonous function, therefore the inverse function $\psi^{-1}(x)$ is uniquely determined for $0 \leqq x \leqq 1$. Hence from (49) we get

$$\Gamma^{-1}(x) = \psi^{-1}(x) - \lambda(1-x),$$

i. e. $\Gamma^{-1}(x)$ is uniquely determined for $0 \leqq x \leqq 1$. Now we investigate whether $\Gamma(s)$ can also be determined uniquely.

$\Gamma^{-1}(0)=\infty$ and $\Gamma^{-1}(1)=0$ and for $0<x<1$ the derivative

$$\frac{d\Gamma^{-1}(x)}{dx} = \lambda + \frac{1}{\psi'(\psi^{-1}(x))}$$

exists and is a monotonically increasing function. In $x=1$ this derivative has the value $\lambda - \frac{1}{\alpha}$. If $\lambda\alpha \leqq 1$, then this derivative is not positive and

$\Gamma^{-1}(x)=0$ has no root in the interval $0 \leqq x < 1$, but $\Gamma^{-1}(1)=0$. If $\lambda\alpha > 1$, then this derivative is negative and $\Gamma^{-1}(x)=0$ has one and clearly only one root in the interval $0 \leqq x < 1$. Let $x = p^*$ denote the smallest positive root of $\Gamma^{-1}(x)=0$. In the interval $(0 \leqq x \leqq p^*)$ the function $\Gamma^{-1}(x)$ is a monotonically decreasing continuous function; therefore its inverse function $\Gamma(s)$ exists, it is the Laplace transform of the sought $G(x)$ for $0 \leqq s < \infty$. This function $\Gamma(s)$ is uniquely determined by the prescribed conditions. The functional equation has in general more solutions but only one of them satisfies the mentioned conditions. Now $\Gamma(0)=p^*$, that is $\lim\limits_{x\to\infty} G(x) = p^*$. Further the roots of the equation $\Gamma^{-1}(x)=0$ agree with those of the equation $\psi(\lambda(1-x))=x$. This completes the proof of the theorem.

EXAMPLE. If $H(x)=1-e^{-x/\alpha}$ $(x \geqq 0)$, then $\psi(s)=1/(1+\alpha s)$ and hence

$$\Gamma(s)=\frac{1+\alpha(s+\lambda)-\sqrt{[1+\alpha(s+\lambda)]^2-4\lambda\alpha}}{2\lambda\alpha}. \tag{51}$$

Now (49) has two solutions corresponding to the positive and negative sign of the root, respectively. We must take the solution with negative square root because of the condition $\Gamma(\infty)=0$. Inverting (51), we obtain

$$G'(x)=e^{-\frac{(1+\lambda\alpha)x}{\alpha}} I_1\left(\frac{2\sqrt{\lambda\alpha x}}{\alpha}\right)\frac{1}{\sqrt{\lambda\alpha x}} \tag{52}$$

where $I_1(x)=J_1(ix)/i$ and $J_1(x)$ denotes the Bessel function of genus 1. Here $G(\infty)=1$ if $\lambda\alpha \leqq 1$, while $G(\infty)=1/\lambda\alpha$ if $\lambda\alpha > 1$.

(49) can be used excellently for the determination of the moments of $G(x)$.

Now we shall determine the average service periode

$$\mu=\int_0^\infty x\,dG(x) \tag{53}$$

directly from (45). From (45) we obtain

$$\mu=\alpha+\sum_{n=1}^{\infty} n\mu\int_0^\infty e^{-\lambda y}\frac{(\lambda y)^n}{n!}\,dH(y). \tag{54}$$

This can be seen directly too: $\mu$ is composed of the average service time $\alpha$ of the first person, and if $n$ persons arrive during the service time of the first — the probability whereof is

$$\int_0^\infty e^{-\lambda y}\frac{(\lambda y)^n}{n!}\,dH(y)$$

— then, by the discussion given above, the average service time of these and the others coming in the meanwhile is equal to $n\mu$.

Performing the summation in (54), we have

(55) $$\mu = \alpha + \lambda\alpha\mu.$$

*If* $\lambda\alpha < 1$, *then*

(56) $$\mu = \frac{\alpha}{1-\lambda\alpha},$$

*however, if* $\lambda\alpha \geqq 1$, *then* $\mu = \infty$.

We have mentioned that in case $\lambda\alpha > 1$ the probability that the service period is finite equals $p^* = \Gamma(0)$. Then the question arises: what is the conditional expected value of the service period under the assumption that the service time is finite? This conditional expected value is

(57) $$\mu^* = -\frac{\Gamma'(0)}{\Gamma(0)}.$$

Here $\Gamma(0) = p^*$ and by (49) $\Gamma'(0) = \psi'(\lambda(1-p^*))[1-\lambda\Gamma'(0)]$, from which $\Gamma'(0)$ can be determined. Finally, we obtain

(58) $$\mu^* = \frac{-\psi'(\lambda(1-p^*))}{p^*[1+\lambda\psi'(\lambda(1-p^*))]}.$$

In case of an exponentially distributed service time $p^* = 1/\lambda\alpha$ if $\lambda\alpha > 1$ and

(59) $$\mu^* = \frac{\alpha}{\lambda\alpha - 1}.$$

In case of a constant service time ($\chi_n = \alpha$), $p^*$ is the smallest positive root of the equation $\lambda\alpha p^* e^{-\lambda\alpha p^*} = \lambda\alpha e^{-\lambda\alpha}$ and

(60) $$\mu^* = \frac{\alpha}{1-\lambda\alpha p^*}.$$

Concerning the service period an other problem arises: How many persons will be served in a service period? Let $f_j$ denote the probability that a service periode consists of $j$ ($j = 1, 2, \ldots$) services. The probability that during the service of a single person $n$ persons arrive at the service counter, is given by

(61) $$p_n = \int_0^\infty e^{-\lambda x}\frac{(\lambda x)^n}{n!}\, dH(x) \qquad (n = 0, 1, 2, \ldots),$$

as mentioned above. Now it is clear that

(62) $$f_j = \sum_{\substack{n_1+n_2+\ldots+n_j = j-1 \\ n_1+n_2+\ldots+n_k \geqq k \ (k=1,2,\ldots,j-1)}} p_{n_1} p_{n_2} \ldots p_{n_j}.$$

However, it seems more convenient to determine the generating function

(63) $$F(\omega) = \sum_{j=1}^{\infty} f_j \omega^j.$$

We have the recurrence formulae

$$(64)\qquad \begin{cases} f_1 = p_0, \\ f_j = \sum_{n=1}^{j-1} p_n \sum_{j_1+j_2+\dots+j_n=j-1} f_{j_1} f_{j_2} \cdots f_{j_n}. \end{cases}$$

The reasoning is perfectly analogous to that used to establish formula (45) for $G(x)$. Passing to the generating function from (64) and using

$$(65)\qquad \psi(\lambda(1-\omega)) = \sum_{n=0}^{\infty} p_n \omega^n,$$

we obtain the following functional equation for $F(\omega)$:

$$(66)\qquad F(\omega) = \omega\psi[\lambda - \lambda F(\omega)].$$

THEOREM 7. *The generating function $F(\omega)$ of the probabilities $f_j$ is the uniquely determined, for $|\omega| \leqq 1$ analytic solution of the functional equation*

$$(67)\qquad F(\omega) = \omega\psi[\lambda - \lambda F(\omega)]$$

*subject to the condition $F(0) = 0$. Let $x = p^*$ be the smallest positive real root of the equation $\psi(\lambda(1-x)) = x$; then $\lim_{\omega\to 1} F(\omega) = p^*$.*

PROOF. It is clear that the probabilities $f_j$ exist. Hence the generating function $F(\omega)$ exists also for $|\omega| \leqq 1$, $F(0) = 0$, $F(\omega)$ is an analytic function for $|\omega| \leqq 1$ and $F(\omega)$ satisfies the functional equation (67). Knowing $F(\omega)$, the probabilities $\{f_j\}$ can be determined uniquely. Therefore, it is sufficient to show that the conditions mentioned determine $F(\omega)$ uniquely. It is sufficient to determine $F(\omega)$ for real values with $0 \leqq \omega \leqq 1$, for then $F(\omega)$ can be determined for all $\omega$ with $|\omega| \leqq 1$. Let $x = F(\omega)$ in (67), then $\omega = F^{-1}(x)$ and by (67) we have

$$F^{-1}(x) = \frac{x}{\psi(\lambda(1-x))}.$$

Now $F^{-1}(0) = 0$, $F^{-1}(1) = 1$ and $\dfrac{dF^{-1}(x)}{dx}$ is monotonically decreasing for $0 \leqq x \leqq 1$. The value of the derivative in $x = 1$ is $1 - \lambda\alpha$. If this derivative is non-negative, that is $\lambda\alpha \leqq 1$, then $F^{-1}(x)$ is monotonically increasing for $0 \leqq x \leqq 1$ and $F^{-1}(1) = 1$. If this derivative is positive, that is $\lambda\alpha > 1$, then $F^{-1}(x) = 1$ has one and only one root $x = p^*$ in the interval $0 < x < 1$; consequently, $F^{-1}(x)$ is monotonically increasing for $0 \leqq x \leqq p^*$ and $F^{-1}(p^*) = 1$. Hence, if $x = p^*$ denotes the smallest positive root of $F^{-1}(x) = 1$, then $F^{-1}(x)$ can be uniquely inverted for $0 \leqq x \leqq p^*$. $F(0) = 0$, $F(\omega)$ is a continuous function of $\omega$ and $\lim_{\omega\to 1} F(\omega) = p^*$. By the above procedure $F(\omega)$ is determined uniquely. $p^*$ denotes the probability that the service period consists of a finite number of services.

EXAMPLE. Let $\chi_n = \alpha$ (constant). Then $\psi(s) = e^{-s\alpha}$ and it is clear that

(68) $$G(x) = \sum_{j=1}^{\left[\frac{x}{\alpha}\right]} f_j.$$

So the determination of $G(x)$ is reduced to that of the probabilities $f_j$. Now $F(\omega)$ satisfies the following functional equation:

(69) $$F(\omega) = \omega e^{-\lambda\alpha[1-F(\omega)]}.$$

If we use the notations $y = \lambda\alpha F(\omega)$ and $x = \omega e^{-\lambda\alpha}\lambda\alpha$, then we obtain the relation

(70) $$y e^{-y} = x.$$

Hence we have to determine the functions $y = y(x)$, i. e. the inverse function of (70). As it is well known, the branch passing through the origin of $y = y(x)$ may be represented as the power series

(71) $$y(x) = \sum_{j=1}^{\infty} \frac{j^{j-1}}{j!} x^j$$

for $|x| \leq 1/e$, i. e. for the values $|\omega| \leq 1$ we obtain

(72) $$F(\omega) = \sum_{j=1}^{\infty} \frac{e^{-\lambda\alpha j}(\lambda\alpha j)^{j-1}}{j!} \omega^j.$$

We remark that from (69) we obtain the differential equation

$$\omega F'(\omega) - F(\omega) = \lambda\alpha\omega F(\omega) F'(\omega)$$

whose solution corresponding to the initial condition $F(0) = 0$ is given also by (72).

So, by (72), we obtain

(73) $$G(x) = \sum_{j=1}^{\left[\frac{x}{\alpha}\right]} \frac{e^{-\lambda\alpha j}(\lambda\alpha j)^{j-1}}{j!}.$$

By (62) we may put

(74) $$f_j = e^{-\lambda\alpha j}(\lambda\alpha)^{j-1} \sum_{\substack{n_1+n_2+\dots+n_j = j-1 \\ n_1+n_2+\dots+n_k \geq k \ (k=1,2,\dots,j-1)}} \frac{1}{n_1!\, n_2! \dots n_j!} = \frac{e^{-\lambda\alpha j}(\lambda\alpha j)^{j-1}}{j!},$$

hence, (73) may be obtained in this way too.

If $\chi_n$ is exponentially distributed, then

$$F(\omega) = \frac{(1+\lambda\alpha) - \sqrt{(1+\lambda\alpha)^2 - 4\lambda\alpha\omega}}{2\lambda\alpha}$$

from which

$$f_j = \frac{1}{2}\binom{2j}{j} \frac{(\lambda\alpha)^{j-1}}{(1+\lambda\alpha)^{2j-1}}.$$

2. *Determination of* $F(t, 0)$. In our paper [16] we have dealt with stochastic processes in which pauses and happenings with distribution function $1-e^{-\lambda t}$ and $G(x)$, respectively, are alternating. We have shown that if the process starts in the moment $t=0$ with a pause, then the probability that in the interval $(t, t+\Delta t)$ a happening (here: service period) begins is $f(t)\Delta t+o(\Delta t)$ where $f(t)$ is continuous and is defined uniquely by the integral equation of Volterra type

(75) $$f(t)=\lambda-\lambda\int_0^t f(t-x)\,[1-G(x)]\,dx.$$

Now we have $f(t)\Delta t+o(\Delta t)=F(t,0)\lambda\Delta t+o(\Delta t)$. In fact, in the interval $(t, t+\Delta t)$ a service period begins when in the moment $t$ there is no waiter, the probability whereof is $F(t, 0)$, and in the interval $(t, t+\Delta t)$ somebody arrives at the service, the probability whereof is $\lambda\Delta t+o(\Delta t)$. Hence

(76) $$F(t, 0)=\frac{f(t)}{\lambda}.$$

By (75), for $\Re(s)>0$ we have

(77) $$\int_0^\infty e^{-st}F(t, 0)\,dt=\frac{1}{s+\lambda-\Gamma(s)}.$$

Hence $F(t, 0)$ may be determined uniquely ($F(0, 0)=1$).

In our paper [16] we have proved that the limit

(78) $$\lim_{t\to\infty} f(t)=\frac{\lambda}{1-\lambda\mu}$$

exists, if $\mu$ is finite; $\mu$ is the mean value of $G(x)$. Hence by (76) and (55) we have

(79) $$\lim_{t\to\infty} F(t, 0)=F^*(0)=\frac{1}{1+\lambda\mu}=1-\lambda\alpha \quad \text{for} \quad \lambda\alpha<1.$$

So we have proved the existence of the limit $F^*(0)$ in the case homogeneous in time. It was used in the preceding chapter.

3. *Determination of the probability of the state* $E_j$. Let us consider the homogeneous process in the stationary case. Let $P_j$ ($j=0, 1, 2, \ldots$) denote the probability that, in a moment chosen arbitrarily, the system has the state $E_j$ ($j=0, 1, 2, \ldots$), i. e. for $t\to\infty$, the number of persons waiting and being served is $j$. $P_0$ is the probability that the system is in the state $E_0$, i. e. there is no person waiting or being served. As we have seen, in the case $\lambda\alpha<1$ we have

(80) $$P_0=1-\lambda\alpha.$$

If $\lambda\alpha\geqq 1$, then, as we have proved, $\mu=\infty$ and by the theory of recurrent events we have $P_0=0$. (See W. FELLER [6], [7].)

In his work [9] A. JA. KHINTCHINE has shown that in the stationary state (i. e. for $t \to \infty$) the probability that the system has the state $E_j$ at the end of a service time, agrees with the probability that in an arbitrary moment the system has the state $E_j$. Thus it suffices to establish these probabilities for the endpoints of the service times only. The random variables, $\zeta_1, \zeta_2, \zeta_3, \ldots$ describe the states at the endpoints of the consecutive service times, where $\zeta_n = j$ if at the end of the $n$-th service the system has the state $E_j$. The process described by the random variables $\{\zeta_n\}$ forms a Markov chain with the transition probabilities

$$P(\zeta_n = k+j-1 \mid \zeta_{n-1} = k) = p_j = \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^j}{j!} \, dH(x) \tag{81}$$

for $j = 0, 1, 2, \ldots$ and $k \geqq 1$, and

$$P(\zeta_n = j \mid \zeta_{n-1} = 0) = p_j \tag{82}$$

for $k = 0$.

It may be easily proved that in our case we can apply a well-known theorem (Cf. W. FELLER [7], p. 325): If the states of an irreducible chain are ergodic, then, independently of the initial value $\zeta_1$, the limiting probabilities

$$\lim_{n \to \infty} P(\zeta_n = j) = P_j \qquad (j = 0; 1, 2, \ldots) \tag{83}$$

exist and the values $P_j$ may be determined uniquely from the system of equations

$$P_j = P_{j+1} p_0 + P_j p_1 + \cdots + P_1 p_j + P_0 p_j, \qquad (j = 0, 1, 2, \ldots). \tag{84}$$

In our case the Markov chain is irreducible, for each state may be reached from any other state in a sufficient number of steps. As known, all the states of an irreducible chain belong to the same class. In case $\lambda \alpha < 1$, as we have seen, the state $E_0$ is ergodic, hence the same holds for every state. If $\lambda \alpha \geqq 1$, then the state $E_0$ is a recurrent-null state, hence so are all other states, i. e. $P_j = 0$ for every $j$.

Let

$$\Pi(\omega) = \sum_{j=0}^{\infty} P_j \omega^j \tag{85}$$

and

$$\pi(\omega) = \sum_{j=0}^{\infty} p_j \omega^j = \int_0^\infty e^{-\lambda(1-\omega)x} \, dH(x) = \psi(\lambda(1-\omega)); \tag{86}$$

then by making use of these generating functions, from (79) we obtain

$$\omega \Pi(\omega) = \pi(\omega) [\Pi(\omega) + P_0 \omega - P_0]. \tag{87}$$

Hence

$$\Pi(\omega)=P_0\frac{\pi(\omega)}{1-\dfrac{1-\pi(\omega)}{1-\omega}}. \tag{88}$$

If $\omega\to 0$, then

$$\frac{1-\pi(\omega)}{1-\omega}\to\lambda\alpha, \tag{89}$$

and hence $\Pi(0)=P_0/(1-\lambda\alpha)$. Considering that $\Pi(0)=1$ or $P_0=1-\lambda\alpha$ if $\lambda\alpha<1$, we obtain

$$\Pi(\omega)=\frac{(1-\lambda\alpha)\psi(\lambda(1-\omega))}{1-\dfrac{1-\psi(\lambda(1-\omega))}{1-\omega}}. \tag{90}$$

EXAMPLE. 1) Let $\chi_n$ be an exponentially distributed random variable, i. e. $H(x)=1-e^{-x/\alpha}$. Then $\psi(s)=1/(1+\alpha s)$ and we have

$$\Pi(\omega)=\frac{1-\lambda\alpha}{1-\lambda\alpha\omega}. \tag{91}$$

Hence

$$P_j=(1-\lambda\alpha)(\lambda\alpha)^j. \tag{92}$$

2) $\chi_n\equiv\alpha$ (constant). Then $\psi(s)=e^{-s\alpha}$ and

$$\Pi(\omega)=\frac{(1-\lambda\alpha)e^{-\lambda\alpha(1-\omega)}}{1-\dfrac{1-e^{-\lambda\alpha(1-\omega)}}{1-\omega}}. \tag{93}$$

Hence

$$P_j=(1-\lambda\alpha)\sum_{k=1}^{j,j-1}(-1)^{j-k}e^{k\lambda\alpha}\left(\frac{(k\lambda\alpha)^{j-k}}{(j-k)!}+\frac{(k\lambda\alpha)^{j-k-1}}{(j-k-1)!}\right). \tag{94}$$

## § 4. Waiting time problems in case of a non-Poisson process

Now, concerning the moments $t_n$ of the arrivals of the persons at the service, let us suppose that the consecutive time differences $\tau_n=t_n-t_{n-1}$ $(n=1,2,\ldots;\ t_0=0)$ are equidistributed mutually independent positive random variables with the common distribution function $P(\tau_n\leqq x)=Q(x)$. In the case of a Poisson process homogeneous in time we had $Q(x)=1-e^{-\lambda x}(x\geqq 0)$. Now we shall deal with the more general case when $Q(x)$ is arbitrary. In the sequel we also suppose that the service times $\chi_1,\chi_2,\chi_3,\ldots$ are equidistributed mutually independent random variables with the common distribution function $H(x)$.

In Chapter 2 we have already mentioned that the random function $\eta(t)$ so defined is in general not a Markov process but the moments $t_n$ are Markov points of the process. Let $\eta(t_n-0)=\eta_n$, i. e. the waiting time of the

$n$-th arriving person, and $\eta(0)=\eta_0$. According to what has been said, the sequence $\{\eta_n\}$ of random variables is a Markov chain. Knowing $\eta_0$, the random variables $\eta_n$ may be determined step by step from the following equation:

$$\eta_{n+1}=\begin{cases}\eta_n+\chi_n-\tau_{n+1} & \text{if} \quad \tau_{n+1}-\chi_n<\eta_n,\\ 0 & \text{if} \quad \tau_{n+1}-\chi_n\geqq\eta_n.\end{cases} \tag{95}$$

Let $P(\eta_n\leqq x)=F_n(x)$ denote the distribution function of the random variable $\eta_n$. Starting from $F_0(x)$, the sequence $\{F_n(x)\}$ may be determined step by step by means of the recurrence formula

$$F_{n+1}(x)=\int_0^\infty K(x,y)\,dF_n(y) \qquad (n=0,1,2,\ldots) \tag{96}$$

where

$$K(x,y)=\int_0^\infty [1-Q(u+y-x)]\,dH(u). \tag{97}$$

In fact, the transition probabilities of the Markov process described by the sequence $\{\eta_n\}$ of variables are $P(\eta_{n+1}\leqq x\,|\,\eta_n=y)=K(x,y)$, so that (96) holds. The computations may be perfected conveniently by the following two-step recurrence formulae:

$$K_n(x)=\int_0^\infty H(x-y)\,dF_n(y) \tag{98}$$

and

$$F_{n+1}(x)=\int_0^\infty [1-Q(y-x)]\,dK_n(y). \tag{99}$$

A further question is: When does the limiting distribution function $\lim_{n\to\infty} F_n(x)=F(x)$ exist and how can $F(x)$ be determined? Here we quote the work by D. V. LINDLEY [14] in which the following theorem is proved:

*For the existence of the limiting distribution function* $\lim_{n\to\infty} F_n(x)=F(x)$ *it is necessary and sufficient that either* $M(\chi_n)\leqq M(\tau_n)$ *or* $\tau_n-\chi_n\equiv 0$ *holds. If* $M(\chi_n)\geqq M(\tau_n)$ *and* $\tau_n-\chi_n\not\equiv 0$, *then* $\lim_{n\to\infty} F_n(x)=0$ *for every x. The limiting distribution* $F(x)$, *if it exists, is independent of the initial distribution* $F_0(x)$ *and is the unique solution of the integral equation*

$$F(x)=\int_0^\infty K(x,y)\,dF(y). \tag{100}$$

In his investigations LINDLEY starts from equation (95) and proves that it characterizes a random walk problem which can be discussed by making use of the laws of large numbers and recurrent events.

The theorem stated above may be proved by general theorems concerning Markov chains or it may be reduced to the proof of the ergodicity of the state $E_0$ with a finite recurrence time if $\alpha < \Theta$ where $\alpha = M(\chi_n)$ and $\Theta = M(\tau_n)$.

It is also interesting to investigate the following problem. If the process $\eta(t)$ is defined for all $t$ as it was described in Chapter 2, then determine conditions for the existence of the limiting distribution function $\lim\limits_{t\to\infty} F(t,x) = F^*(x)$ and a method to obtain it. Here we state the following

THEOREM 8. *If $0 < \alpha < \Theta < \infty$ and $Q(x)$ does not agree with a distribution function which has only jumps at some integral multiplies of a given number, then there exists the limiting distribution function $\lim\limits_{t\to\infty} F(t,x) = F^*(x)$, independently of the initial distribution $F_0(x)$ and may be given by*

$$F^*(x) = \frac{1}{\Theta}\int_0^x [1-Q(y)]K(y+x)\,dy, \tag{101}$$

*where*

$$K(x) = \int_0^\infty F(x-y)\,dH(y) \tag{102}$$

*and $F(x)$ is the unique solution of the integral equation* (100).

For the proof let us introduce the new variable $\zeta(t)$ which denotes the distance between the moment $t$ and the immediately preceding moment $t_n$ $(n=0,1,2,\ldots)$. The process $\zeta(t)$ so defined essentially agrees with a special case of the renewal process investigated by J. L. DOOB [3]. Under our conditions we may apply Theorem 12 of DOOB. According to this theorem, the limiting distribution function

$$Q^*(x) = \lim_{t\to\infty} P(\zeta(t) \leq x) = \frac{1}{\Theta}\int_0^x [1-Q(u)]\,du \tag{103}$$

exists.

If $t\to\infty$ and $t_n$ denotes the moment of arrival of the last person arriving in the interval $(0,t)$, then with probability one we have $n\to\infty$ and $\lim\limits_{n\to\infty} P(\eta(t_n) \leq x) = K(x)$, where $K(x)$ is given by (102). So it is clear that

$$F^*(x) = \int_0^x K(x+y)\,dQ^*(y) \tag{104}$$

which is the same as (101).

The uniqueness as well as the independence of the initial distribution $F_0(x)$ follow from the theorem stated for $F(x)$.

We remark that if $F^*(x)$ will be defined as the distribution function of $\eta(t)$ in the moment $t=t^*$ chosen "at random" in a process of an infinitely long time, then it suffices to suppose regarding $Q(x)$ only that $0<\Theta<\infty$, where $\Theta$ is its mean value.

EXAMPLE. Apart from the waiting times here we may also quote as example the filling and evacuating of cisterns. This example is taken from the book of G. P. BOEV [1] (p. 311, example 170). Let $\eta(t)$ denote the provision of the cistern in the moment $t$. Cistern are filled with periodes of a year. If we suppose that the quantity of leaving water is the same in the unit time until the evacuating follows, then we have a phenomenon which may be described by the above model. In the stationary case the distribution function of the provisions in the cistern will be given by $F^*(x)$ in (101).

If we suppose that the water leaving in the unit time is proportional to the provisions, then the model described in our paper [17] may be applied.

INSTITUTE FOR APPLIED MATHEMATICS
OF THE HUNGARIAN ACADEMY OF SCIENCES

*(Received 1 December 1952)*

## Bibliography

[1] Г. П. Боев, Теория Вероятностей (Москва, 1950).
[2] G. DOETSCH, *Handbuch der Laplace Transformation*, Bd. 1 (Basel, 1949).
[3] J. L. DOOB, Renewal theory from the point of view of the theory of probability, *Transactions of the American Mathematical Society*, **63** (1948), pp. 422—438.
[4] W. FELLER, Zur Theorie der stochastischen Prozesse, *Mathematische Annalen*, **113** (1936), pp. 13—60.
[5] W. FELLER, On theory of stochastic processes with particular reference to applications, *Proceedings of the First Berkeley Symposium*, (1949), pp. 403—432.
[6] W. FELLER, Fluctuation theory of recurrent events, *Transactions of the American Mathematical Society*, **67** (1949), pp. 98—119.
[7] W. FELLER, *An introduction to probability theory and its applications* (New-York, 1950).
[8] TH. C. FRY, *Probability and its engineering uses* (New-York, 1928), pp. 372—387.
[9] А. Я. Хинчин, Математическая теория стационарной очереди, Математический Сборник, **39** (1932), pp. 73—84.
[10] D. G. KENDALL, Some problems in the theory of queues, *Journal of the Royal Statistical Society*, Ser. B, **13** (1951), pp. 151—185.
[11] A. N. KOLMOGOROV, Sur le problème d'attente, *Recueil Mathématique*, Математический Сборник, **38** (1931), pp. 101—106.
[12] A. N. KOLMOGOROV—B. V. GNEDENKO, *Független valószínűségi változók összegeinek határeloszlásai* (Budapest, 1951).
[13] D. V. LINDLEY, The theory of queues with a single counter, *Proceedings of the Cambridge Philosophical Society*, **48** (1952), pp. 277—289.

[14] F. Pollaczek, Über eine Aufgabe der Wahrscheinlichkeitsrechnung, *Mathematische Zeitschrift*, **32** (1950), pp. 64—100; II, pp. 729—850.

[15] F. Pollaczek, Répartition des délais d'attente des avions arrivant à un aéroport qui possède *s* pistes d'atterissages, *Comptes Rendus Acad. Sci. Paris*, **232** (1951), p. 1901.

[16] L. Takács, Occurrence and coincidence phenomena in case of happenings with arbitrary distribution law of duration, *Acta Math. Acad. Sci. Hung.*, **2** (1951), pp. 275—298.

[17] L. Takács, On stochastic processes connected with some physical recording apparatuses, *Acta Math. Acad. Sci. Hung.*, **6** (1955) (to be appear).

[18] A. Zygmund, A remark on characteristic functions, *Proceedings of the Second Berkeley Symposium*, (1951), pp. 369 372.

## ИССЛЕДОВАНИЕ ПРОБЛЕМ ВРЕМЕНИ ОЖИДАНИЯ С ПОМОЩЬЮ МАРКОВСКИХ ПРОЦЕССОВ

Л. Такач (Будапешт)

(Резюме)

Лица прибывающие в моментах времени $t_1, t_2, \ldots, t_n, \ldots$ соответственно, обслуживаются в порядке очереди единственным приказчиком. Пусть времена обслуживания $\chi_1, \chi_2, \ldots, \chi_n, \ldots$ следующих друг за другом лиц являются независимыми положительными случайными величинами с одном и том же функцией распределения $H(x)$. Пусть $\Re(s) \geqq 0$

$$\psi(s) = \int_0^\infty e^{-sx}\, dH(x)$$

и $M(\chi_n) = \alpha$.

Обозначим через $\eta(t)$ время ожидания лица при условии что он прибывает в момент, и через $\eta_n$ время ожидания лица прибывающего $n$-ым. Пусть далее $P(\eta(t) \leqq x) = F(t, x)$ и $P(\eta_n \leqq x) = F_n(x)$.

1. Пусть последовательность $\{t_n\}$ моментов прибытия образует процесс Пуассона с плотностью событий $\lambda(t)$. Тогда $\eta(t)$ является марковским процессом и $F(t, x)$ удовлетворяет следующему интегро-дифференциальному уравнению:

$$\frac{\partial F(t,x)}{\partial t} = \frac{\partial F(t,x)}{\partial x} - \lambda(t)\left[F(t,x) - \int_0^x H(x-y)\, d_y F(t,y)\right].$$

С помощью этого уравнения можно однозначно определять $F(t, x)$, зная $F(t, 0)$. (Нетрудно определять $F(t, 0)$, если $\lambda(t) \equiv \lambda$ (постоянна). Пусть $\lim\limits_{t\to\infty} \lambda(t) = \lambda$; тогда $\lim\limits_{t\to\infty} F(t,0) = 1-\lambda\alpha$ если $\lambda\alpha < 1$, и $\lim\limits_{t\to\infty} F(t,0) = 0$ если $\lambda\alpha \geqq 1$. В случае $\lambda\alpha < 1$ существует предельное распределение $\lim\limits_{t\to\infty} F(t,x) = F^*(x)$, и для преобразованная Лапласа—Стильтьеса от $F^*(x)$ имеем

$$\int_0^\infty e^{-sx}\, dF^*(x) = \frac{1-\alpha}{1-\lambda\dfrac{1-\psi(s)}{s}},$$

в соответствии решением заданным раньше Хинчиным для стационарного случая.

2. Пусть последовательность $\{t_n\}$ моментов прибытия образует процесс Пуассона с (постоянной) плотностью событий $\lambda(t) \equiv \lambda$. Тогда последовательность переменных $\{\eta_n\}$ образует однородную цепь Маркова, а функции распределения $F_n(x)$ определяются с помощью следующей рекуррентной формулы:

$$F_n(x) = \int_0^\infty K(x, y)\, dF_{n-1}(y) \qquad (n = 2, 3, \ldots),$$

где

$$K(x, y) = H(x-y) + \int_{x-y}^\infty e^{-\lambda(u+y-x)}\, dH(u).$$

Если $n \to \infty$, то $F_n(x) \to F^*(x)$.

Пусть $G(x)$ — функция распределения периода обслуживания, и пусть

$$\Gamma(s) = \int_0^\infty e^{-sx}\, dG(x).$$

Тогда $\Gamma(s)$ является решением функционального уравнения

$$\Gamma(s) = \psi(s + \lambda - \lambda\Gamma(s)),$$

удовлетворяющим условию $\Gamma(\infty) = 0$ и аналитическим для $\Re(s) \geqq 0$. Пусть $p^*$ — наименьшее положительное число, для которого $p^* = \psi(\lambda(1-p^*))$, тогда $\lim_{x\to\infty} G(x) = p^*$. Если $\lambda\alpha \leqq 1$, то $p^* = 1$, а если $\lambda\alpha > 1$, то $0 < p^* < 1$.

Пусть $f_j$ вероятность того, что период обслуживания состоит из $j$ отдельных актов обслуживания, и пусть $F(\omega) = \sum_{j=1}^\infty f_j \omega^j$. Тогда $F(\omega)$ является решением функционального уравнения

$$F(\omega) = \omega\psi(\lambda - \lambda F(\omega)),$$

удовлетворяющим условию $F(0) = 0$ и аналитическим для $|\omega| \leqq 1$. Далее, имеет место соотношение $\lim_{\omega\to 1} F(\omega) = p^*$, где значение $p^*$ то же как и прежде.

Пусть $P_j$ — вероятность того, что при $t \to \infty$ число ожидающих равняется $j$. Пусть $\Pi(\omega) = \sum_{j=0}^\infty P_j \omega^j$; тогда для $\lambda\alpha < 1$ имеет место

$$\Pi(\omega) = \frac{(1-\lambda\alpha)\,\psi(\lambda(1-\omega))}{1 - \dfrac{1-\psi(\lambda(1-\omega))}{1-\omega}}.$$

3. На конце статьи делаются некоторые замечания относительно случая, когда разности $t_n - t_{n-1}$ одинаково распределенные независимые случайные величины.

# LÖSUNG DER VEKTOR-FUNKTIONALGLEICHUNG DER HOMOGENEN UND INHOMOGENEN $n$-DIMENSIONALEN EINPARAMETRIGEN „TRANSLATION", DER ERZEUGENDEN FUNKTION VON KETTENREAKTIONEN UND DES STATIONÄREN UND NICHT-STATIONÄREN BEWEGUNGSINTEGRALS

Von

J. ACZÉL (Debrecen)

(*Vorgelegt von* A. RÉNYI)

## § 1

Die folgende Vektor-Funktionalgleichung der $n$-dimensionalen einparametrigen (homogenen) Translation (s. z. B. [1][1]), d. h. der $n$-dimensionalen Transformationen mit einem additiven Parameter (auf die — vgl. § 4 — unter gewissen Voraussetzungen durch Einführung eines additiven Parameters jede einparametrige Transformation zurückgeführt werden kann), ist zugleich die Funktionalgleichung der erzeugenden Funktion von Kettenreaktionen im homogenen Fall [2] und auch die Integrale der Vektor-Differentialgleichung (d. h. des Differentialgleichungssystems) einer stationären Bewegung erfüllen diese Gleichung (s. z. B. [3]):

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, t), u] = \mathbf{f}(\mathbf{x}, t+u). \tag{1}$$

(Übrigens folgen auch die Iterierten $\mathbf{f}_v(\mathbf{x}) = \mathbf{f}(\mathbf{x}, v)$ einer Vektorfunktion dieser Gleichung, vgl. § 4.)

Die bezügliche Gleichung für die erzeugende Funktion der inhomogenen Kettenreaktionen [2] lautet

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, s, t), t, u] = \mathbf{f}(\mathbf{x}, s, u). \tag{2}$$

Man sieht auch leicht, daß die Integrale des Differentialgleichungssystems (der Vektor-Differentialgleichung) einer nicht-stationären Bewegung

$$\frac{d}{dt}\mathbf{f}(\mathbf{x}, s, t) = \mathbf{v}(\mathbf{f}, t)$$

[1] Eckige Klammern verweisen auf das Literaturverzeichnis am Ende der Arbeit.

der Gleichung (2) genüge leisten. Es scheint auch zweckmäßig zu sein, als Verallgemeinerung der gewöhnlichen Translationen die der Gleichung (2) genügenden Zuordnungen „inhomogene Translationen“ zu nennen.

In dem § 3 werden wir Folgerungen bezüglich der oben vorgeführten Probleme (z. B. die Zurückführbarkeit der einparametrigen Transformationen auf koordinaten-parallele Verschiebungen, usw.) aus der in dem § 2 vollbrachten Lösung der Funktionalgleichungen (1), (2) ziehen.

Im eindimensionalen Fall wurde die Gleichung (1) und die ähnliche Gleichung für nicht-additiven Parameter schon in den Arbeiten [4], [5] gelöst.

Die Voraussetzungen über die Natur der Funktionen $\mathbf{f}$ werden wir so wählen, daß die Beweise möglichst einfach und klar werden, deshalb werden wir nicht immer die schwächsten Bedingungen suchen. Es ist aber bemerkenswert, daß wir die Gültigkeit unserer Funktionalgleichungen nur auf einer $(n-1)$-dimensionalen Hyperebene voraussetzen und hieraus ihre Gültigkeit schon im ganzen Raum folgt.

## § 2

Bezüglich der Gleichung (1) beweisen wir den

Satz 1. *Die Vektorfunktionen*

$$\mathbf{f}(\mathbf{x}, t) = \mathbf{g}[\mathbf{h}(\mathbf{x}) + t], \tag{3}$$

*wo* $\mathbf{h}(\mathbf{x})$ *eine beliebige ein-eindeutige Vektorfunktion mit der Umkehrfunktion* $\mathbf{g}$ *ist, also*

$$\mathbf{h}[\mathbf{g}(\mathbf{y})] = \mathbf{y}, \tag{4}$$

*und die Addition* $\mathbf{y} + t$ *zwischen Vektoren und Skalaren als Addition des Skalars zur ersten Komponenten des Vektors:*

$$\mathbf{y} + t = \{y_1, y_2, \ldots, y_n\} + t = \{y_1 + t, y_2, \ldots, y_n\}$$

*definiert wird, erfüllen alle die Gleichung*

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, t), u] = \mathbf{f}(\mathbf{x}, t + u). \tag{1}$$

*Wird umgekehrt* (1) *auf der Koordinatenhyperebene*[2] $\mathbf{x} = \mathbf{x}_a = \{a, x_2, \ldots, x_n\}$ *(a konstant) vorausgesetzt, und gibt es ferner für jedes* $\mathbf{x}$ *ein*

$$\mathbf{x}_t = \{t, x_2, \ldots, x_n\} = \mathbf{h}(\mathbf{x}) \tag{5}$$

*derart, daß*

$$\mathbf{g}(\mathbf{x}_t) = \mathbf{f}(\mathbf{x}_a, t) = \mathbf{x} \tag{6}$$

*ist, so folgt*

$$\mathbf{f}(\mathbf{x}, t) = \mathbf{g}[\mathbf{h}(\mathbf{x}) + t] \tag{3}$$

*(und die Gültigkeit der Gleichung* (1)*) für jedes* $\mathbf{x}$ (und $t, u$).

[2] Natürlich kann statt $x_1$ auch $x_2$ usw. konstant gehalten werden, wenn dies die Erfüllung der übrigen Voraussetzungen besser sichert.

Die oben stehende Definition der Addition von Vektoren und Skalaren mag etwas seltsam scheinen, sie leistet aber hier guten Dienst (vgl. auch § 3). Auch die übliche Addition von komplexen Zahlen bzw. Quaternionen mit Skalaren (reellen Zahlen) entspricht dieser Vorschrift.

Beweis. Daß die Funktion (3) die Gleichung (1) erfüllt, sieht man unmittelbar:

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, t), u] = \mathbf{g}[\mathbf{h}(\mathbf{f}(\mathbf{x}, t)) + u] = \mathbf{g}(\mathbf{h}[\mathbf{g}(\mathbf{h}(\mathbf{x}) + t)] + u) =$$
$$= \mathbf{g}[\mathbf{h}(\mathbf{x}) + t + u] = \mathbf{f}(\mathbf{x}, t + u),$$

wo auch auf (4) Rücksicht genommen wurde.

Gilt umgekehrt (1) für $\mathbf{x} = \mathbf{x}_a$ (und alle $t, u$):

$$\mathbf{f}[\mathbf{f}(\mathbf{x}_a, t), u] = \mathbf{f}(\mathbf{x}_a, t + u),$$

dann ist wegen (5)—(6) und der Definition von $\mathbf{x}_t + u = \mathbf{x}_{t+u}$:

$$\mathbf{f}[\mathbf{g}(\mathbf{x}_t), \mathbf{u}] = \mathbf{g}(\mathbf{x}_t + u),$$

also

$$\mathbf{f}(\mathbf{x}, u) = \mathbf{g}[\mathbf{h}(\mathbf{x}) + u]$$

für jedes $\mathbf{x}$ gültig, w. z. b. w.

Der entsprechende Satz für die Gleichung (2) lautet:

Satz 2. *Die Funktionen*

(7) $$\mathbf{f}(\mathbf{x}, t, u) = \mathbf{g}[\mathbf{h}(\mathbf{x}, t), u]$$

*(wo* $\mathbf{y} = \mathbf{h}(\mathbf{x}, t)$ *eine beliebige Vektorfunktion mit dem Parameter* $t$ *ist, die sich bezüglich* $x$ *umkehren läßt:* $\mathbf{x} = \mathbf{g}(\mathbf{y}, t)$, *also*

(8) $$\mathbf{h}[\mathbf{g}(\mathbf{y}, t), t] = \mathbf{y}$$

*gilt) erfüllen immer die Gleichung*

(2) $$\mathbf{f}[\mathbf{f}(\mathbf{x}, s, t), t, u] = \mathbf{f}(\mathbf{x}, s, u).$$

*Wird umgekehrt* (2) *auf der Koordinatenhyperebene* $\mathbf{x} = \mathbf{x}_a$ *erfüllt und gibt es für jedes* $\mathbf{x}$ *und* $t$ *ein*

(9) $$\mathbf{x}_s = \{s, x_2, \ldots, x_n\} = \mathbf{h}(\mathbf{x}, t)$$

*derart, daß*

(10) $$\mathbf{g}(\mathbf{x}_s, t) = \mathbf{f}(\mathbf{x}_a, s, t) = \mathbf{x}$$

*ist, so gilt*

(7) $$\mathbf{f}(\mathbf{x}, t, u) = \mathbf{g}[\mathbf{h}(\mathbf{x}, t), u]$$

*(und deshalb auch die Gleichung* (2)*) für jedes* $\mathbf{x}$ (und $t, u, s$).

Es ist bemerkenswert, daß hier gar keine neue Operation zwischen Vektoren und Skalaren definiert werden mußte.

BEWEIS. Die Funktionen der Gestalt (7) erfüllen offenbar die Gleichung (2):

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, s, t), t, u] = \mathbf{g}(\mathbf{h}[\mathbf{f}(\mathbf{x}, s, t), t], u) = \mathbf{g}[\mathbf{h}(\mathbf{g}[\mathbf{h}(\mathbf{x}, s), t], t), u] = \\ = \mathbf{g}[\mathbf{h}(\mathbf{x}, s), u] = \mathbf{f}(\mathbf{x}, s, u)$$

mit Rücksicht auf (8).

Ist umgekehrt (2) für $\mathbf{x} = \mathbf{x}_a$ (und alle $s, t, u$) erfüllt:

$$\mathbf{f}[\mathbf{f}(\mathbf{x}_a, s, t), t, u] = \mathbf{f}(\mathbf{x}_a, s, u),$$

so gilt wegen (9)—(10)

$$\mathbf{f}[\mathbf{g}(\mathbf{x}_s, t), t, u] = \mathbf{g}(\mathbf{x}_s, u),$$

also

$$\mathbf{f}(\mathbf{x}, t, u) = \mathbf{g}[\mathbf{h}(\mathbf{x}, t), u]$$

für jedes $\mathbf{x}$, w. z. b. w.

Die Frage nach der Eindeutigkeit der Darstellungen (7) und (3) läßt sich folgenderweise erledigen:

SATZ 3. *Ist*

(11) $$\mathbf{g}[\mathbf{h}(\mathbf{x}, t), u] = \bar{\mathbf{g}}[\bar{\mathbf{h}}(\mathbf{x}, t), u],$$

*wo* $\mathbf{h}$ *bzw.* $\bar{\mathbf{h}}$ *die einzigen Umkehrfunktionen von* $\mathbf{g}$ *bzw.* $\bar{\mathbf{g}}$ *sind, so gilt*

(12) $$\bar{\mathbf{g}}(\mathbf{y}, u) = \mathbf{g}[\mathbf{k}(\mathbf{y}), u]$$

$$(\bar{\mathbf{h}}(\mathbf{x}, u) = \mathbf{l}[\mathbf{h}(\mathbf{x}, u)], \quad \textit{wo} \quad \mathbf{l} = \mathbf{k}^{-1}, \quad \textit{d. h.} \quad \mathbf{k}[\mathbf{l}(\mathbf{x})] = \mathbf{x})$$

*und diese erfüllen auch* (11). *D. h. in* (7) *kann* $\mathbf{g}$ *durch die in* (12) *gegebenen* $\bar{\mathbf{g}}$ *und nur durch diese ersetzt werden, also ist die Funktion* $\mathbf{g}$ *in* (7) *nur bis auf eine n-dimensionale Vektorfunktion* $\mathbf{k}$ *bestimmt.*

BEWEIS. Aus (11) folgt, wenn wir $\bar{\mathbf{h}}(\mathbf{x}, t) = \mathbf{y}$, $\mathbf{x} = \bar{\mathbf{g}}(\mathbf{y}, t)$ schreiben mit $t = t_0$:

$$\bar{\mathbf{g}}(\mathbf{y}, u) = \mathbf{g}(\mathbf{h}[\bar{\mathbf{g}}(\mathbf{y}, t_0), t_0], u) = \mathbf{g}[\mathbf{k}(\mathbf{y}), u].$$

Hieraus folgt auch

$$\bar{\mathbf{h}}(\mathbf{x}, u) = \mathbf{l}[\mathbf{h}(\mathbf{x}, u)],$$

und diese Funktionen erfüllen auch wirklich (11):

$$\bar{\mathbf{g}}[\bar{\mathbf{h}}(\mathbf{x}, t), u] = \mathbf{g}(\mathbf{k}[\bar{\mathbf{h}}(\mathbf{x}, t)], u) = \mathbf{g}[\mathbf{k}(\mathbf{l}[\mathbf{h}(\mathbf{x}, t)]), u] = \mathbf{g}[\mathbf{h}(\mathbf{x}, t), u],$$

w. z. b. w.

Im homogenen Fall gilt ähnlich der

SATZ 4. *Aus*

(13) $$\mathbf{g}[\mathbf{h}(\mathbf{x}) + t] = \bar{\mathbf{g}}[\bar{\mathbf{h}}(\mathbf{x}) + t]$$

*folgt*

(14) $$\bar{\mathbf{g}}(\mathbf{y}) = \mathbf{g}[\varkappa(\mathbf{y}^{(n-1)}) + y_1],$$

*(wo die Bezeichnungen* $\mathbf{y} = \{y_1, y_2, \ldots, y_n\}$, $\mathbf{y}^{(n-1)} = \{y_2, \ldots, y_n\}$ *verwendet wurden), und diese Funktion erfüllt mit ihrer Umkehrfunktion* $\bar{\mathbf{h}}(\mathbf{x})$ *die Gleichung*

(13) *tatsächlich. D. h. in* (3) *kann* $\mathbf{g}$ *durch die in* (14) *gegebenen* $\bar{\mathbf{g}}$ *und nur durch diese ersetzt werden, also ist die Funktion* $\mathbf{g}$ *in* (3) *nur bis auf eine n-dimensionale Vektorfunktion eines* $(n-1)$*-dimensionalen Vektors (d. h. n Funktionen von* $(n-1)$ *Veränderlichen) bestimmt.*

Der BEWEIS dieses Satzes ist merkwürdigerweise mehr verwickelt als der des Satzes 3. (14) folgt aus (13) folgenderweise: Wir bezeichnen in (13) $\bar{\mathbf{h}}(\mathbf{x})+t=\mathbf{y}$, $\mathbf{x}=\mathbf{g}(\mathbf{y}-t)$, dann gilt $\bar{\mathbf{g}}(\mathbf{y})=\mathbf{g}(\mathbf{h}[\mathbf{g}(\mathbf{y}-t)]+t)$, wo $t$ noch beliebig gewählt werden kann. Wählen wir $t=y_1$, so erhalten wir tatsächlich

$$\bar{\mathbf{g}}(\mathbf{y})=\mathbf{g}(\mathbf{h}[\bar{\mathbf{g}}(\mathbf{y}^{(n-1)})]+y_1)=\mathbf{g}[\varkappa(\mathbf{y}^{(n-1)})+y_1].$$

Aber schon die Konstruktion der Umkehrfunktion $\bar{\mathbf{h}}(\mathbf{x})$ von $\mathbf{g}(\mathbf{y})$ ist nicht mehr trivial:

Aus

$$\mathbf{x}=\bar{\mathbf{g}}(\mathbf{y})=\mathbf{g}[\varkappa(\mathbf{y}^{(n-1)})+y_1]=\mathbf{g}\left[\begin{pmatrix}\varkappa_1(\mathbf{y}^{(n-1)})+y_1\\ \varkappa^{(n-1)}(\mathbf{y}^{(n-1)})\end{pmatrix}\right],\quad \mathbf{y}=\bar{\mathbf{h}}(\mathbf{x})$$

folgt

$$\varkappa(\mathbf{y}^{(n-1)})+y_1=\mathbf{h}(\mathbf{x}),$$

d. h.

$$\varkappa_1(\mathbf{y}^{(n-1)})+y_1=h_1(\mathbf{x}),$$
$$\varkappa^{(n-1)}(\mathbf{y}^{(n-1)})=\mathbf{h}^{(n-1)}(\mathbf{x})$$

und falls wir die Umkehrfunktion von $\mathbf{x}^{(n-1)}=\varkappa^{(n-1)}(\mathbf{y}^{(n-1)})$ mit $\mathbf{y}^{(n-1)}=$ $=\lambda^{(n-1)}(\mathbf{x}^{(n-1)})$ bezeichnen, so wird

$$\begin{cases}\bar{\mathbf{h}}^{(n-1)}(\mathbf{x})=\mathbf{y}^{(n-1)}=\lambda^{(n-1)}[\mathbf{h}^{(n-1)}(\mathbf{x})],\\ \bar{h}_1(\mathbf{x})=y_1=h_1(\mathbf{x})-\varkappa_1(\mathbf{y}^{(n-1)})=h_1(\mathbf{x})-\varkappa_1(\lambda^{(n-1)}[\mathbf{h}^{(n-1)}(\mathbf{x})]\end{cases}$$

die gesuchte Umkehrfunktion von (14) sein. Nun erfüllt sie aber zusammen mit (14) die Gleichung (13) tatsächlich

$$\bar{\mathbf{g}}[\bar{\mathbf{h}}(\mathbf{x})+t]=\bar{\mathbf{g}}\left[\begin{pmatrix}\bar{h}_1(\mathbf{x})+t\\ \bar{\mathbf{h}}^{(n-1)}(\mathbf{x})\end{pmatrix}\right]=\bar{\mathbf{g}}\left[\begin{pmatrix}h_1(\mathbf{x})-\varkappa_1(\mathbf{y}^{(n-1)})+t\\ \lambda^{(n-1)}[\mathbf{h}^{(n-1)}(\mathbf{x})]\end{pmatrix}\right]=$$
$$=\mathbf{g}\left[\begin{pmatrix}\varkappa_1(\lambda^{(n-1)}[\mathbf{h}^{(n-1)}(\mathbf{x})])+h_1(\mathbf{x})-\varkappa_1(\lambda^{(n-1)}[\mathbf{h}^{(n-1)}(\mathbf{x})])+t\\ \varkappa^{(n-1)}(\lambda^{(n-1)}[\mathbf{h}^{(n-1)}(\mathbf{x})])\end{pmatrix}\right]=$$
$$=\mathbf{g}\left[\begin{pmatrix}h_1(\mathbf{x})+t\\ \mathbf{h}^{(n-1)}(\mathbf{x})\end{pmatrix}\right]=\mathbf{g}[\mathbf{h}(\mathbf{x})+t],$$

(da $\varkappa^{(n-1)}[\lambda^{(n-1)}(\mathbf{x}^{(n-1)})]=\mathbf{x}^{(n-1)}$ ist), w. z. b. w.

## § 3

Die Lösungen (3) bzw. (7) der Gleichungen (1) bzw. (2) bedeuten, daß *mittels der Koordinatentransformation*

$$\bar{\mathbf{x}}=\mathbf{h}(\mathbf{x}),\quad \bar{\mathbf{y}}=\mathbf{h}(\mathbf{y}),$$

*bzw. der Zuordnung des beweglichen Punktes*

$$\bar{\mathbf{x}}(t)=\mathbf{h}(\mathbf{x},t)$$

*zum Punkte* $\mathbf{x}$ *(und ähnlich* $\bar{\mathbf{y}}(t) = \mathbf{h}(\mathbf{y}, t)$ *zu* $\mathbf{y}$) *die Transformationen mit einem additiven Parameter*

$$\mathbf{y} = \mathbf{f}(\mathbf{x}, t),$$

*bzw. die „inhomogenen Translationen“*

$$\mathbf{y} = \mathbf{f}(\mathbf{x}, t, u)$$

*sich in die zur* $x_1$ *Achse parallelen Verschiebung*

$$\bar{\mathbf{y}} = \bar{\mathbf{x}} + t,$$

*bzw. in die Transformation*

$$\bar{\mathbf{y}}(u) = \bar{\mathbf{x}}(t) \tag{15}$$

*überführen lassen, falls die Transformationen* $\mathbf{f}(\mathbf{x}, t)$ *bzw.* $\mathbf{f}(\mathbf{x}, t, u)$ *auf der Koordinatenhyperebene* $x_1 = a$ *eine Schar bilden, d. h. den Gleichungen* (1) *bzw.* (2) *genügen und dort auch die Auflösbarkeitsbedingungen* (5)—(6) *bzw.* (9)—(10) *erfüllt sind (und dann bleiben die Schareigenschaften* (1) *bzw.* (2) *im ganzen Raum erhalten).*

(15) läßt sich so deuten, daß der zu $\mathbf{x}$ geordnete bewegliche Punkt im Zeitpunkt $u$ eben dort ist, wo der zu $\mathbf{y} = \mathbf{f}(\mathbf{x}, t, u)$ geordnete im Zeitpunkt $t$. Der homogene Fall schließt sich hierin durch bewegliche Punkte mit axenparallelen Bahnen ein.

Bei unserer Deutung gibt der Satz 4 aus einer geeigneten Koordinatentransformation alle solche an und Ähnliches gilt auch vom Satz 3.

Für die wahrscheinlichkeitstheoretische Anwendung unserer Gleichungen ist die Deutung unserer Resultate weniger an der Hand liegend. Jedenfalls kann bemerkt werden, daß die erzeugenden Funktionen der Wahrscheinlichkeiten bei Kettenreaktionen dem Werte-$n$-tuppel $(1, 1, \ldots, 1)$ immer den Funktionenwert $(1, 1, \ldots, 1)$ zuordnen, deshalb wird bei den oben beschriebenen Zuordnungen dem Einheitspunkt immer ein bezüglich der neuen Transformation fixer Punkt entsprechen. Ähnliches gilt oft auch bezüglich des Anfangspunktes $(0, 0, \ldots, 0)$.

Wir wenden uns endlich den stationären und nicht-stationären Bewegungen zu. Wie bekannt ([3]), erfüllen die Ortsvektoren eines beweglichen Punktes $\mathbf{f}(\mathbf{x}, t)$ mit dem Anfangspunkt $\mathbf{x}$, die der Differentialgleichung der stationären Bewegung

$$\frac{d\mathbf{f}(\mathbf{x}, t)}{dt} = \mathbf{v}[\mathbf{f}(\mathbf{x}, t)] \tag{16}$$

genügen, die Gleichung (1). Ähnlich erfüllen die Ortsvektoren des beweglichen Punktes $\mathbf{f}(\mathbf{x}, s, t)$, der sich zur Ausgangszeit $s$ am Ort $\mathbf{x}$ befindet, falls die Bewegung nichtstationär ist (also der Gleichung

$$\frac{d\mathbf{f}(\mathbf{x}, s, t)}{dt} = \mathbf{v}[\mathbf{f}(\mathbf{x}, s, t), t] \tag{17}$$

genügt), die Gleichung (2). Den Zusammenhang dieser Funktionen $\mathbf{v}$ mit unseren Vektorfunktionen geben die folgenden zwei Sätze an.

SATZ 5. *Die Funktionen* $\mathbf{v}(\mathbf{x})$ *in der Differentialgleichung der stationären Bewegung*

$$\frac{d\mathbf{f}(\mathbf{x},t)}{dt}=\mathbf{v}[\mathbf{f}(\mathbf{x},t)], \tag{16}$$

*(deren Lösungen* $\mathbf{f}(\mathbf{x},t)$ *der Gleichung* (1) *genüge leisten), lassen sich (mittels* $\mathbf{f}$ *bzw.) mittels der im Ausdruck* (3) *von* $\mathbf{f}$ *figurierenden Funktion* $\mathbf{h}$ *folgenderweise ausdrücken:*

$$\mathbf{v}(\mathbf{x})=\left.\frac{d\mathbf{f}(\mathbf{x},t)}{dt}\right|_{t=0}=\mathbf{H}'(\mathbf{x})^{-1}\mathbf{e},$$

*wo* $\mathbf{e}$ *der Vektor* $e_1=1, e_2=\cdots=e_n=0$ *ist und* $\mathbf{H}'(\mathbf{x})$ *die als regulär vorausgesetzte derivierte Matrix der Vektorfunktion* $\mathbf{h}(\mathbf{x})$ *und* $\mathbf{H}'(\mathbf{x})^{-1}$ *ihre Inverse ist.*

BEWEIS. Aus (3) folgt, wie wir schon sahen,

$$\mathbf{h}[\mathbf{f}(\mathbf{x},t)]=\mathbf{h}(\mathbf{x})+t.$$

Derivieren wir diese Gleichung nach $t$, so können wir

$$\mathbf{H}'[\mathbf{f}(\mathbf{x},t)]\frac{d\mathbf{f}(\mathbf{x},t)}{dt}=\mathbf{e} \tag{18}$$

schreiben. Aus (3) (oder auch aus (1)) folgt aber

$$\mathbf{f}(\mathbf{x},0)=\mathbf{x}$$

und aus (16) erhalten wir für $t=0$

$$\mathbf{v}(\mathbf{x})=\left.\frac{d\mathbf{f}(\mathbf{x},t)}{dt}\right|_{t=0},$$

folglich wird aus (18) für $t=0$

$$\mathbf{v}(\mathbf{x})=\mathbf{H}'(\mathbf{x})^{-1}\mathbf{e},$$

w. z. b. w.

SATZ 6. *Die Funktionen* $\mathbf{v}(\mathbf{x},t)$ *von* (17) *stehen mit den Funktionen* $\mathbf{f}(\mathbf{x},s,t)$ *und den Funktionen* $\mathbf{h}(\mathbf{x},t)$ *von* (7) *im folgenden Zusammenhang*

$$\mathbf{v}(\mathbf{x},t)=\left.\frac{d\mathbf{f}(\mathbf{x},s,t)}{dt}\right|_{s=t}=-\mathbf{H}'(\mathbf{x},t)^{-1}\frac{d\mathbf{h}(\mathbf{x},t)}{dt},$$

*wo* $\mathbf{H}'$ *wieder die derivierte Matrix von* $\mathbf{h}(\mathbf{x},t)$ *bezüglich* $\mathbf{x}$ *bezeichnet.*

BEWEIS. Aus (7) (oder (2)) folgt

$$\mathbf{f}(\mathbf{x},t,t)=\mathbf{x}, \tag{19}$$

und aus

$$\mathbf{v}[\mathbf{f}(\mathbf{x},s,t),t]=\frac{d\mathbf{f}(\mathbf{x},s,t)}{dt} \tag{17}$$

für $s=t$

$$\mathbf{v}(\mathbf{x},t)=\left.\frac{d\mathbf{f}(\mathbf{x},s,t)}{dt}\right|_{s=t}. \tag{20}$$

Andererseits derivieren wir wieder die mit (7) äquivalente Gleichung

$$\mathbf{h}[\mathbf{f}(\mathbf{x}, s, t), t] = \mathbf{h}(\mathbf{x}, s)$$

nach $t$

$$\mathbf{H}'[\mathbf{f}(\mathbf{x}, s, t), t]\frac{d\mathbf{f}(\mathbf{x}, s, t)}{dt} + \frac{d\mathbf{h}(\mathbf{y}, t)}{dt}\bigg|_{\mathbf{y}=\mathbf{f}(\mathbf{x}, s, t)} = 0$$

und substituieren $s$ $t$, dann wird mit Rücksicht auf (19) und (20)

$$\mathbf{v}(\mathbf{x}, t) = -\mathbf{H}'(\mathbf{x}, t)^{-1}\frac{d\mathbf{h}(\mathbf{x}, t)}{dt}.$$

Dies und (20) ergeben die Behauptung des Satzes 6.

Die Sätze 6 bzw. 5 lassen sich für die Komponenten folgenderweise aussprechen:

$$v_k = -\frac{\Delta_k}{\Delta}, \quad \textit{wo} \quad \Delta = \frac{\partial(h_1, \ldots, h_n)}{\partial(x_1, \ldots, x_n)} \quad \textit{und} \quad \Delta_k = \frac{\partial(h_1, \ldots, h_{k-1}, h_k, h_{k+1}, \ldots, h_n)}{\partial(x_1, \ldots, x_{k-1}, t, x_{k+1}, \ldots, x_n)}$$

*bzw.*

$$\Delta_k = \frac{\partial(h_1 - t, h_2, \ldots, h_{k-1}, h_k, h_{k+1}, \ldots, h_n)}{\partial(x_1, x_2, \ldots, x_{k-1}, t, x_{k+1}, \ldots, x_n)}$$

*ist.*

## § 4

Die bisherigen Interpretationen unserer Gleichungen liefen parallel im homogenen und inhomogenen Fall. Wir schließen mit zwei Bemerkungen die sich nur auf die homogene Gleichung (1) beziehen.

Erstens sei bemerkt, daß auch *die iterierten* $\mathbf{f}_t(\mathbf{x}) = \mathbf{f}(\mathbf{x}, t)$ *einer Vektorfunktion* $\mathbf{f}(\mathbf{x})$, d. h. eines Funktionen-n-tuppels $f_1(x_1, \ldots, x_n), f_2(x_1, \ldots, x_n), \ldots, f_n(x_1, x_2, \ldots x_n)$, *der Gleichung* (1) *genügen*, da man von den Iterierten die Erfüllung der Relation

$$\mathbf{f}_u[\mathbf{f}_t(\mathbf{x})] = \mathbf{f}_{t+u}(\mathbf{x})$$

zu fordern pflegt.

Zweitens sei darauf hingewiesen wie sich die Transformationen mit einem nicht-additiven Parameter auf den Fall mit additivem Parameter, d. h. auf die „homogenen Translationen" der Gleichung (1) zurückführen lassen.

SATZ 7. *Falls die n-dimensionalen, einparametrigen Transformationen $f(x, t)$ mit einem (nicht-additiven) Parameter eine Schar in dem Sinne*

(21) $$\mathbf{f}[\mathbf{f}(\mathbf{x}, t), u] = \mathbf{f}(\mathbf{x}, t \circ u)$$

*bilden, [wo $t \circ u$ eine stetige rechts kürzbare Operation (skalare Funktion) von $t$ und $u$ ist], und $\mathbf{f}(\mathbf{x}, t)$ effektiv ist, d. h. aus $\mathbf{f}(\mathbf{x}, t) \equiv \mathbf{f}(\mathbf{x}, u)$ $t = u$ folgt, so kann ein additiver Parameter eingeführt werden, d. h. es existiert eine (skalare) stetige und streng monotone Funktion $\varphi(t)$ derart, daß die neue Transformation*

(22) $$\bar{\mathbf{f}}(\mathbf{x}, \mathrm{t}) = \mathbf{f}[\mathbf{x}, \varphi(t)]$$

*schon die Gleichung*

$$\bar{\mathbf{f}}[\bar{\mathbf{f}}(\mathbf{x}, t), u] = \bar{\mathbf{f}}(\mathbf{x}, t+u)$$

*der homogenen Translation erfüllt.*

BEWEIS. Aus (21) folgt

$$\mathbf{f}[\mathbf{x}; (s \circ t) \circ u] = \mathbf{f}[\mathbf{f}(\mathbf{x}, s \circ t), u] = \mathbf{f}(\mathbf{f}[\mathbf{f}(\mathbf{x}, s), t], u) = \\ = \mathbf{f}[\mathbf{f}(\mathbf{x}, s), t \circ u] = \mathbf{f}[\mathbf{x}, s \circ (t \circ u)],$$

was wegen der Effektivität von $\mathbf{f}$ die *Assoziativität*

$$(s \circ t) \circ u = s \circ (t \circ u)$$

von $t \circ u$ bedeutet. Es ist ebenfalls aus der Effektivität zu sehen, daß aus $t \circ u_1 = t \circ u_2$ wegen

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, t), u_1] = \mathbf{f}(\mathbf{x}, t \circ u_1) = \mathbf{f}(\mathbf{x}, t \circ u_2) = \mathbf{f}[\mathbf{f}(\mathbf{x}, t), u_2]$$

$u_1 = u_2$ folgt, also $t \circ u$ auch von *links kürzbar* ist. Da nun, wie wir gleich beweisen werden, jede stetige assoziative und beiderseitig kürzbare Operation (Funktion) $t \circ u$ in der Gestalt

(23) $$t \circ u = \varphi[\varphi^{-1}(t) + \varphi^{-1}(u)]$$

darstellbar ist, folgt aus (21)

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, t), u] = \mathbf{f}(\mathbf{x}, \varphi[\varphi^{-1}(t) + \varphi^{-1}(u)]),$$

also mit der Bezeichnung (22) und $\varphi^{-1}(t) = v$, $\varphi^{-1}(u) = w$

$$\bar{\mathbf{f}}[\mathbf{f}(\mathbf{x}, v), w] = \bar{\mathbf{f}}(\mathbf{x}, v+w),$$

w. z. b. w.

Die noch übrig gebliebene Verifikation von (23) folgt aus dem

SATZ 8. *Ist die Operation $t \circ u$ zwischen zwei reellen Zahlen $t, u$ eines Intervalles stetig, assoziativ und beiderseitig kürzbar, so gibt es eine stetige, streng monotone Funktion $\varphi(t)$ derart, daß*

(23) $$t \circ u = \varphi[\varphi^{-1}(t) + \varphi^{-1}(u)],$$

*($\varphi^{-1}$ ist die inverse Funktion von $\varphi$).*

Der BEWEIS benutzt den Satz (s. [6]), daß für jede stetige *streng monoton wachsende* assoziative Operation (23) gilt. Wir müssen also noch beweisen, daß *aus unseren Voraussetzungen dvs streng monotone Wachsen von $t \circ u$ in beiden Veränderlichen folgt.* Die strenge Monotonität selbst folgt natürlich aus der Stetigkeit und beiderseitigen Kürzbarkeit von $t \circ u$, wir müssen aber die Möglichkeit des streng monotonen Abnehmens ausschließen. Gäbe es z. B. ein $u_0$ derart, daß für jedes Paar $t_1 < t_2$

(24) $$t_1 \circ u_0 - t_2 \circ u_0 > 0$$

wäre, so hat jedes $u$ dieselbe Eigenschaft, denn falls es noch ein $u$ gebe für das

$$t_1 \circ u - t_2 \circ u < 0,$$

so gäbe es wegen der Stetigkeit auch ein $u_1$ zwischen $u_0$ und $u$ mit

$$t_1 \circ u_1 - t_2 \circ u_1 = 0,$$

was der rechtsseitigen Kürzbarkeit widersprechen würde. Gibt es also ein $u_0$, derart, daß für jedes Paar $t_1 < t_2$ (24) gilt, so folgt für jedes $u$ aus $t_1 < t_2$

$$t_1 \circ u > t_2 \circ u$$

und

$$t_2 \circ (u \circ v) < t_1 \circ (u \circ v) = (t_1 \circ u) \circ v < (t_2 \circ u) \circ v = t_2 \circ (u \circ v),$$

was unmöglich ist. Ähnlich kann $t \circ u$ auch in der zweiten Veränderlichen nicht abnehmen. Damit ist der Beweis vollendet.

*(Eingegangen am 4. Mai 1953.)*

**Bemerkung bei der Korrektur.** (Am 9. Juni 1955.) Ein dem Satz 2 analoger Satz für $n = 1$ ist in der eleganten Arbeit von S. GOLAB, Sur les objets géométriques non différentiels, *Bulletin de l'Acad. Polon. des Sci. et des Lettres Classe des Sci. Math. et Nat.* A, (1949), S. 67—72 zu finden. Dies zeigt zugleich, daß (2) auch in der Theorie der geometrischen Objekten eine Anwendung hat. Auch (1) kann in dieser Theorie angewendet werden. Auf diese Frage wollen wir noch zurückkehren.

## Literaturverzeichnis

[1] G. VRANCEANU, *Leçons de géométrie différentielle I* (Bucarest, 1947), S. 65—77.

[2] Б. А. Севастьянов, Теория ветвящихся случайных процессов, Успехи Мат. Наук, **6** (1951), в. 6 (46), S. 47—99, insb. 54.

[3] В. В. Степанов, Курс дифференциальных уравнений (Москва—Ленинград, 1950), S. 267—270.

[4] J. ACZÉL, L. KALMÁR, J. G. MIKUSINSKI, Sur l'équation de translation, *Studia Math.*, **12** (1951), S. 112—116.

[5] J. ACZÉL, Über einparametrige Transformationen, *Publ. Math. (Debrecen)*, **1** (1950), S. 243—247.

[6] J. ACZÉL, Sur les opérations définies pour nombres réels, *Bulletin de la Soc. Math. de France*, **76** (1948), S. 59—64.

# РЕШЕНИЕ ВЕКТОРНОГО ФУНКЦИОНАЛЬНОГО УРАВНЕНИЯ ОДНОРОДНЫХ И НЕОДНОРОДНЫХ $n$-МЕРНЫХ ОДНОПАРАМЕТРИЧЕСКИХ „ТРАНСЛАЦИЙ", В СВЯЗИ С ПРОИЗВОДЯЩЕЙ ФУНКЦИЕЙ ЦЕПНЫХ РЕАКЦИЙ И ИНТЕГРАЛОМ СТАЦИОНАРНЫХ И НЕСТАЦИОНАРНЫХ ДВИЖЕНИЙ

Я. Ацел (Дебрецен)

(Резюме)

Целью настоящей работы является доказать, что функции

$$\mathbf{f}(\mathbf{x}, t) = \mathbf{g}[\mathbf{h}(\mathbf{x}) + t] \qquad (\mathbf{y} + t = \{y_1, y_2, \ldots, y_n\} + t = \{y_1 + t, y_2, \ldots, y_n\})$$

и

$$\mathbf{f}(\mathbf{x}, t, u) = \mathbf{g}[\mathbf{h}(\mathbf{x}, t), u]$$

($\mathbf{h}$ произвольный, $\mathbf{h}[\mathbf{g}(\mathbf{y})] = \mathbf{h}[\mathbf{g}(\mathbf{y}, t), t] = \mathbf{y}$) являются самыми общими решениями функциональных уравнений

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, t), u] = \mathbf{f}(\mathbf{x}, t + u)$$

и

$$\mathbf{f}[\mathbf{f}(\mathbf{x}, s, t), t, u] = \mathbf{f}(\mathbf{x}, s, u)$$

соответственно, если требовать выполненность условий разрешимости (5)—(6) и (9)—(10) соответственно. Эти уравнения имеют применения, указанные в заглавии работы. Излагаются также некоторые дальнейшие обобщения, следствия и применения.

# EXTREMUM PROPERTIES OF THE REGULAR POLYTOPES

By

L. FEJES TÓTH (Budapest)

(*Presented by* G. HAJÓS)

It is more than a century ago that L. SCHLÄFLI[1] discovered the higher dimensional analogues of the five Platonic solids. He found that besides the regular simplex, cross polytope and measure polytope, i. e. the analogues of the regular tetrahedron $\{3,3\}$, octahedron $\{3,4\}$ and the cube $\{4,3\}$, there are only three further convex regular polytopes, all in four dimensions. These are the self dual $\{3,4,3\}$ made up from 24 $\{3,4\}$, 3 around each edge, $\{3,3,5\}$ bounded by 600 $\{3,3\}$, 5 around each edge, and its dual, $\{5,3,3\}$ having 120 $\{5,3\}$-cells (dodecahedra), 3 around each edge.

The great variety of the well-known extremum properties of the regular polygons, and especially of the equilateral triangle, suggests analogous properties of the regular solids and polytopes. In fact, many extremum properties of the regular polygons can easily be transferred to the regular simplex and a part of them to the cross polytope or measure polytope.[2] Also $\{5,3\}$ and $\{3,5\}$ (icosahedron) can be characterised by different extremum postulates.[3] But — though the theory of the regular polytopes has a vast literature[4] — little, if anything, is known in this direction about the "non-trivial" regular polytopes $\{3,4,3\}$, $\{3,3,5\}$ and $\{5,3,3\}$ mentioned above.[5]

In the present paper we shall give a simple proof[6] of the fact that among those convex four-dimensional polytopes of given insphere which are topologically isomorphic with $\{3,3,3\}$ (regular simplex), or $\{3,3,4\}$ (cross

[1] Theorie der vielfachen Kontinuität, *Denkschriften der schweizerischen naturforschenden Gesellschaft*, **38** (1901), pp. 1—237; — *Ges. Abh.* I (Basel, 1950), pp. 167—387.

[2] See e. g. the remarks on the end of the paper.

[3] L. FEJES TÓTH, *Lagerungen in der Ebene, auf der Kugel und im Raum* (Berlin—Göttingen—Heidelberg, 1953).

[4] H. S. M. COXETER, *Regular polytopes* (London, 1948).

[5] L. FEJES TÓTH, On close-packings of spheres in spaces of constant curvature, *Publicationes Math.* (*Debrecen*), **3** (1953), pp. 158—167.

[6] Compare our proof with M. GOLDBERG, The isoperimetric problem for polyhedra, *Tôhoku Math. J.*, **42** (1935), pp. 226—236, and H. HADWIGER, Zur isoperimetrischen Ungleichung für $k$-dimensionale konvexe Polyeder, *Nagoya Math. J.*, 5 (1953), pp. 39—44.

polytope), or $\{3, 3, 5\}$ (600-cell), or $\{3, 4, 3\}$ (24-cell), the corresponding regular polytope has the minimal volume.

It was J. STEINER who first compared polyhedra of the same topologic type, probably in order to characterize all the five regular solids by a unique extremum condition.[7] We can guess that the consideration of isomorphic polytopes has the same effect also in higher dimensions. For example, there is no doubt that the above proposition holds for $\{4, 3, 3\}$ (measure polytope) and $\{5, 3, 3\}$ (120-cell) too, but the proof of this conjecture seems to be not so simple. On the other hand, for the remaining polytopes our proof yields a little more:

*Among all convex four-dimensional polytopes of given insphere bounded by* 5, 16 *or* 600 *tetrahedra or* 24 *polyhedra isomorphic with the regular octahedron the corresponding regular polytope has the minimal volume.*

Let $P$ be in 4-space a convex polytope bounded by $n$ tetrahedra $t_1, \ldots, t_n$ and containing the unit sphere $S$. Let $\tau_i$ be the spherical tetrahedron arising from $t_i$ by central projection upon $S$. The sum of the volumes[8] $\tau_i$ equals the surface (3-dimensional content) of $S$:

$$\sum_{i=1}^{n} \tau_i = 2\pi^2.$$

We assert that

$$t_i \geqq v(\tau_i)$$

where $v(\tau)$ denotes the volume of a regular tetrahedron whose hyperplane (3-space) touches $S$ at the centre of the tetrahedron and the central projection of which has a volume equal to $\tau$. This is equivalent to the fact that among the tetrahedra $t$, of given volume, whose hyperplanes do not intersect $S$, the regular tetrahedron touching $S$ at its centre has maximal projection.

It is easy to show that if $t$ is not entirely contained in a sufficiently large sphere concentric with $S$ the projection $\tau$ of $t$ becomes arbitrarily small. Therefore we can restrict ourself to tetrahedra lying in a fixed sphere and (with respect to the continuity of the functional $\tau$) the existence of an extremal tetrahedron follows from the theorem of WEIERSTRASS.

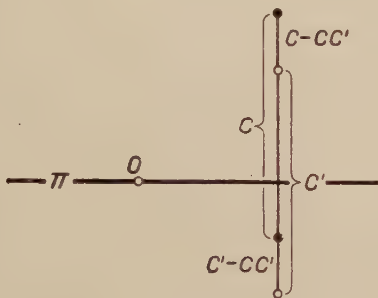
Obviously, we can suppose that the hyperplane of $t$ touches $S$ at a point $O$ since otherwise we could increase the volume of its projection by a translation. Let $\Pi$ be a plane through $O$, orthogonal to an edge of $t$. Let us replace all chords $c$ of $t$ orthogonal to $\Pi$ by the chords $c'$ of the same length and lying on the same line symmetrically with respect to $\Pi$. This process, called symmetrisation of STEINER,[9] carries $t$ into a new tetrahedron $t'$ of the

[7] Such a characterisation is impossible by comparing polyhedra of given number of faces or vertices.

[8] We denote a body and its volume by the same symbol.

[9] The symmetrisation of a tetrahedron and octahedron was already used by STEINER in order to establish the isoperimetric property of $\{3, 3\}$ and $\{3, 4\}$.

same volume $t' = t$ symmetric with respect to $\Pi$. The fact that any inner point of the segment $c' - cc'$ lies nearer to $O$ than any point of $c - cc'$ involves a volume preserving representation of $t$ upon $t'$ which lets the points of $tt'$ invariant and carries the other ones into points lying nearer to $O$. But, since among two volume elements of $t$ of equal volume that one lying nearer to $O$ has a greater projection upon $S$, the projection of $t'$ is greater than or equal to the projection of $t$. Equality holds only if $t$ is symmetric with respect to $\Pi$.




It follows that the best tetrahedron must be symmetric with respect to any plane through $O$ orthogonal to an edge, i. e. it is regular and has $O$ for its centre. This completes the proof of our assertion.

The above remark on the volume elements of $t$ makes it evident that $v(\tau)$ is a convex function of $\tau$ $(0 \leqq \tau < \pi^2)$. Hence, by Jensen's inequality

$$4P \geqq \sum_{i=1}^{n} t_i \geqq \sum_{i=1}^{n} v(\tau_i) \geqq n \cdot v\left(\frac{1}{n}\sum_{i=1}^{n} \tau_i\right) = n \cdot v\left(\frac{2\pi^2}{n}\right).$$

The case of equality is obvious and implies just the desired extremum property of the regular 5-, 16- and 600-cell.

The proof of our theorem concerning $\{3, 4, 3\}$ runs analogously by replacing the word "tetrahedron" by "octahedron" or, more precisely, by "polyhedron isomorphic with the regular octahedron". But in this case we symmetrise first on a plane orthogonal to a diagonal of the octahedron. Under such a symmetrisation the topologic type of the octahedron remains unaltered. Three successive symmetrisations on such planes passing through $O$ yield an octahedron symmetric with respect to $O$ and having mutually orthogonal diagonals. Consequently, we can restrict ourself to the simple case of octahedra of this property. Symmetrising such an octahedron first on a plane through $O$ orthogonal to an edge, then on a plane through $O$ orthogonal to a diagonal radiating from one endpoint of this edge, we obtain an octahedron of the same property. It follows that the best octahedron must be symmetric also with respect to any plane through $O$ orthogonal to an edge, i. e. it must be regular and to have $O$ for its centre.

An analogous proof concerning $\{4, 3, 3\}$ and $\{5, 3, 3\}$ miscarries on the fact that in case of a polyhedron isomorphic with $\{4, 3\}$ or $\{5, 3\}$ the symmetrisation on any plane generally alters the type of the polyhedron.

We close our discussions with some remarks.

The process used in the proof of the inequality $t \geqq v(\tau)$ shows at the same time that among the tetrahedra $t$ of given volume the regular tetrahedron $\bar{t}$ concentric with a sphere $S$ contains the greatest part of the volume of $S$:

$$tS \leqq \bar{t}S.$$

This is a generalisation of the well-known facts that among the tetrahedra inscribed in or circumscribed about to a sphere the regular tetrahedron has the maximal and minimal volume, respectively.

Another consequence of the symmetrisation is the fact that among the tetrahedra the regular one has the least quotient $R/r$ where $R$ denotes the circumradius and $r$ the inradius.[10]

Analogous propositions hold for polyhedra isomorphic with $\{3, 4\}$. Further, the theorem concerning $R/r$ can be transferred by aid of polar reciprocation from $\{3, 4\}$ to $\{4, 3\}$.[11] All these proofs can be extended without difficulty to higher dimensions. So we obtain, for instance: If $R$ and $r$ are the circumradius and inradius of an $n$-dimensional simplex, then

$$R \geqq nr.$$

In case of an $n$-dimensional polytope isomorphic with the cross polytope or measure polytope we have

$$R^2 \geqq nr^2.$$

*(Received 19 February 1954)*

## ОБ ЭКСТРЕМАЛЬНЫХ СВОЙСТВАХ РЕГУЛЯРНЫХ ПОЛИТОПОВ

Л. Фееш Тот (Будапешт)

(Резюме)

Автором доказывается, что из выпуклых политопов четырех измерений, топологически изоморфных регулярному политопу символа $\{3, 3, 3\}$ или $\{3, 3, 4\}$ или $\{3, 3, 5\}$, и содержащих данную сферу, сам регулярный политоп имеет минимальный объем.

[10] A very simple direct proof of this was given by I. Ádám. (See the book quoted in footnote [3], p. 28.)

[11] Cf. the book cited in [3], p. 131.

# ÜBER GEWISSE EIGENSCHAFTEN ORTHOGONALER SYSTEME DER KLASSE $L^2$ UND DIE EIGENFUNKTIONEN STURM—LIOUVILLESCHER DIFFERENTIALGLEICHUNGEN

Von

MIKLÓS MIKOLÁS (Budapest)

(*Vorgelegt von* B. SZ.-NAGY)

## § 1. Einleitung

**1.1.** Vielfach benutzt man diejenige Eigenschaft des *trigonometrischen Orthogonalsystems,* daß man von ihm durch Differentiation und Integration wieder Orthogonalsysteme erhält, welche überdies mit demselben Intervalle (der Länge $2\pi$) verknüpft sind. Ähnlich ist der Sachverhalt, wie man sofort einsieht, z. B. bei den Tschebyscheffschen Polynomen erster und zweiter Art, allgemeiner bei den Jacobischen, gleichwie bei den Laguerreschen und Hermiteschen Polynomen; d. h. in diesen Fällen ist die abgeleitete Folge, bzw. ein Integralsystem der mit der Gewichtsfunktion multiplizierten Folge wieder inbezug auf dasselbe Intervall nebst einer passenden Belegung orthogonal. Die nach den „klassischen" Systemen gebildeten Orthogonalentwicklungen gehen also bei Differentiation und Integration in *Orthogonalreihen* desselben Typus über; m. a. W.: sie behalten ihren Charakter bei den genannten Operationen. Diese Tatsache ist für die Anwendungen ebenso vorteilhaft, wie die entsprechende im Falle der *Potenzreihen.*

**1.2.** Auf Grund der bisherigen Überlegungen hat die Frage ein gewisses Interesse, inwiefern alles dies sich verallgemeinern läßt. — Dementsprechend wollen wir uns in der vorliegenden Arbeit mit den folgenden Problemen beschäftigen:

I. Es sei $\{\varphi_\lambda(x)\}$ ein beliebiges Orthogonalsystem der Klasse $L^2_w(a, b)$, verknüpft mit der nicht negativen Belegungsfunktion $w(x)$.[1] Es seien gewisse

[1] Es kommen nur solche Belegungsfunktionen (Gewichtsfunktionen) in Betracht, welche nur über einer Nullmenge des Grundintervalles $[a, b]$ verschwinden und darin $L$-integrabel sind. — $L^2_w(a, b)$ bezeichnet die Gesamtheit der in $(a, b)$ meßbaren Funktionen $f(x)$, für welche $\int_a^b w(x)f(x)^2\,dx$ existiert. Ferner benutzt man ($w(x)\equiv 1$) die übliche Bezeichnung $L^p(a, b)$ für die Klasse der in $(a, b)$ meßbaren Funktionen $f(x)$ mit dabei $L$-integrablen $|f(x)|^p$. — Hier und durchwegs in dieser Arbeit kommt nur der *Lebesguesche Integralbegriff* zur Anwendung.

Konstanten $c_1, c_2 \ldots$ und eine Funktion $Q(x) \geqq 0$ $(a \leqq x \leqq b)$ angegeben. Gesucht wird eine notwendige und hinreichende Bedingung dafür, daß die Folge $\{\Phi_\lambda(x)+c_\lambda\}$ mit $\Phi_\lambda(x)=\int_a^x w(t)\varphi_\lambda(t)\,dt$ $(a \leqq x \leqq b;\ \lambda=1,2,\ldots)$, bzw. das erweiterte System $1, \{\Phi_\lambda(x)+c_\lambda\}$ inbezug auf $(a, b)$ nebst der Belegungsfunktion $Q(x)$ orthogonal ist.[2]

II. Es sei $\{\varphi_\lambda(x)\}$ ein Orthogonalsystem von in $(a, b)$ differenzierbaren Funktionen für $(a, b)$, verknüpft mit der ebenda nicht negativen Belegungsfunktion $w(x)$. Es sei noch eine Funktion $g(x) \geqq 0$ $(a \leqq x \leqq b)$ angegeben. Gesucht wird eine notwendige und hinreichende Bedingung dafür, daß die Ableitungsfolge $\{\varphi_\lambda'(x)\}$ ein Orthogonalsystem zur Belegung $g(x)$ für das Intervall $(a, b)$ bildet.

Eng schließen sich die weiteren Aufgaben heran:

III. Es ist eine möglicherweise weite Klasse solcher Orthogonalsysteme anzugeben und näher zu untersuchen, welche gleichzeitig die in I und II genannten Eigenschaften besitzen.

IV. Wie lassen sich die „klassischen" Orthogonalsysteme, d. h. die Jacobischen, Laguerreschen, Hermiteschen Polynome und die Sturm—Liouvilleschen Eigenfunktionen auf Grund der in dieser Richtung gefundenen Resultate charakterisieren?

V. Sei eine (allgemeine) Fourierreihe betrachtet, von welcher man durch Differentiation (Integration) wieder eine Orthogonalreihe erhält; was kann über die Konvergenz bzw. Summabilität der ursprünglichen und abgeleiteten (integrierten) Reihe ausgesprochen werden?

**1.3.** Diese Probleme scheinen in der gegebenen Fassung neu zu sein. Was IV betrifft, so sind in den letzten Jahrzehnten mehrere Untersuchungen publiziert worden, welche von den üblichen Definitionen (Rodriguessche Formeln, Orthogonalisierung, Erzeugungsfunktionen, Eigenwertaufgaben) abweichende, eigenartige Kennzeichnung einer oder anderer Klasse von *Orthogonalpolynomen* enthalten.[3] J. EGERVÁRY gab z. B. in einer schönen Arbeit gewisse charakteristische geometrische Eigenschaften der Legendreschen und Tschebyscheffschen Polynome an,[4] während W. HAHN und H. L. KRALL den Satz bewiesen haben: *unter allen orthogonalen Polynomsystemen* besitzen die klassischen im wesentlichen[5] allein die Eigenschaft, daß die aus ihnen durch

[2] Wie wir sehen werden, ist der Fall $c_1=c_2=\cdots=0$ bei den klassischen Orthogonalpolynomen ausgezeichnet.

[3] Vgl. die Bibliographie [28].

[4] Vgl. [5]. — Der Satz rührt von L. FEJÉR her. (Vgl. EGERVÁRY, l. c., S. 18.)

[5] D. h. von trivialen linearen Transformationen abgesehen. — Vgl. [14], [15], [21], [22], [23], ferner [30], [28], [2].

Differentiation (oder wiederholtes Differenzieren) entstehende Polynomfolge gleichfalls nebst einer geeigneten Belegung inbezug auf dasselbe Intervall orthogonal ist. Nach der Entstehung dieser Arbeit ist eine Notiz von J. ACZÉL erschienen, welche sich auf eine vereinfachte Behandlung der klassischen Orthogonalpolynome mittels einer Rodriguesschen Formel bezieht.[6] — In diesem Zusammenhange hat man noch einen ganz speziellen Satz von B. GAGAEFF zu erwähnen,[7] wonach es außer 1, $\cos x$, $\sin x$, $\cos 2x$, $\sin 2x, \ldots$ wesentlich kein anderes Orthogonalsystem gibt, welches sich bei Derivation bis auf Zahlenfaktoren reproduziert.[8]

V umfaßt offenbar die klassische Frage über die gliedweise Differenzierbarkeit und Integrierbarkeit einer gewöhnlichen Fourierreihe; die zweite ist bekanntlich stets bestattet, bezüglich der ersten aber folgt aus dem Fejérschen Grundsatze kurz und elegant, daß die stetige Differenzierbarkeit der Summenfunktion die $C_1$-Summierbarkeit (zur Derivierten) der abgeleiteten Reihe nach sich zieht.[9] Dieses schöne Fejérsche Ergebnis hat etwa zehn Jahre später zu mehrerlei weiteren Untersuchungen über trigonometrische Fourierreihen Anlaß gegeben.

**1. 4.** Die erwähnten Probleme werden wir vor allem im Falle vollständiger Systeme behandeln.

Bezüglich I und II ergibt sich in § 2 die notwendige und hinreichende Bedingung, daß $\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ einer *Integralgleichung* der Form

$$\varphi_\lambda(x) = \varrho_\lambda \int_a^x Q(t)[\Phi_\lambda(t) + c_\lambda]\, dt + K_\lambda \tag{1. 1}$$

mit bestimmten konstanten $\varrho_\lambda$, $c_\lambda$, $K_\lambda$ *fast überall in* $(a, b)$ genügt (vgl. Satz 1),[10] bzw. (was im Falle stetiger $\varphi_\lambda(x)$ auf dasselbe hinausgeht), daß in $(a, b)$ eine Relation vom Typus

$$[g(x)\varphi_\lambda'(x)]' = \varrho_\lambda w(x)\varphi_\lambda(x) \tag{1. 2}$$

nebst Randbedingungen erfüllt ist. (Sätze 2 und 3.)

6 Vgl. [1].

7 Vgl. [10], [12].

8 Beim Lektorieren machte mich Herr Professor B. SZŐKEFALVI-NAGY auf eine mir vorher unzugängliche, neuestens publizierte Arbeit von D. C. LEWIS ([25]) aufmerksam; hierin beweist der Verfasser einige, meinen Sätzen 3 und 4 teilweise entsprechende Behauptungen, seine Resultate stehen aber nur mit meinem Problem II in Berührung. Ich will Prof. SZ.-NAGY für die wertvollen Ratschläge auch an dieser Stelle meinen besonderen Dank ausdrücken. — Wie ich bei der Korrektur bemerkt habe, findet sich ein dem ersten Teile meines Satzes 3 ähnliches, etwas spezielleres Ergebnis schon in einer anderen Arbeit von GAGAEFF ([11]), welche aber von LEWIS nicht zitiert wird.

9 Vgl. [6], [7].

10 Wir geben gleich in 2. 1 eine Verschärfung des Satzes 1, welche sich auf *Lebesgue—Stieltjes-Integrale* bezieht.

M. a. W.: es soll jede Funktion $y=\varphi_\lambda(x)$ $(\lambda=1,2,\ldots)$ eine Lösung der *Sturm—Liouvilleschen Differentialgleichung*

$$[g(x)y']'=\varrho_\lambda w(x)y \tag{1.3}$$

sein.

In § 3 zeigt man u. a. ohne Mühe (vgl. III), daß die, passende Randbedingungen befriedigenden Lösungen einer Gleichung von der Form (1.3) ein die Eigenschaften I und II *gleichzeitig* besitzendes Orthogonalsystem bilden (Satz 4). Die Gesamtheit der betreffenden Folgen $\{\varphi_\lambda(x)\}$ — *„gewöhnlicher SL-Systeme“* — ist somit als eine natürliche Erweiterungsklasse zu betrachten. Es wird auch die naheliegende Frage untersucht, wann die ersten bzw. höheren Ableitungen eines gewöhnlichen *SL*-Systems $\{\varphi_\lambda(x)\}$ gleichfalls zu dieser Klasse gehören (Satz 5); die Bemerkung, daß die Derivierte einer Lösung von (1.3) wieder einer Differentialgleichung vom Sturm—Liouvilleschen Typus genügt, wobei die Koeffizienten einfach mit $g(x)$, $w(x)$ zusammenhängen (Lemma 1), spielt da und später eine wesentliche Rolle.

Die beiden letzten Abschnitte (§§ 4 und 5) enthalten *Anwendungen* bezüglich IV und V.

Sätze 6, 7, 8 liefern je eine solche *Charakterisierung* (in der Gesamtheit aller Orthogonalsysteme der Klasse $L^2$) für das trigonometrische System bzw. für die klassischen Orthogonalpolynome, ferner für beliebige Eigenfunktionen von (1.3) erster oder zweiter Art, welche nur *Eigenschaften allgemeiner Natur*, wie Beschränktheit, Stetigkeit bzw. Differenzierbarkeit, Orthogonalität, Vollständigkeit benutzt.[11]

Satz 9 ergibt endlich eine *Verallgemeinerung des Fejérschen Satzes* über die Differentiation gewöhnlicher Fourierschen Sinusreihen, deren Beweis natürlich weitere Hilfsmittel aus der Theorie der Sturm—Liouvilleschen Randwertprobleme (z. B. asymptotische Formeln, das Äquikonvergenztheorem von Haar) und aus der allgemeinen Theorie der reellen Funktionen erfordert.

## § 2. Untersuchung der Probleme I, II

**2.1.** In den folgenden bezeichnen, wie üblich, $[a, b]$ ein abgeschlossenes, $(a, b)$ ein offenes Intervall, welche endlich oder unendlich sein mögen. Es kommen durchwegs nur reelle Zahlen und Funktionen in Betracht.

Zunächst gilt bezüglich der Frage I der folgende

Satz 1. *Es sei $w(x)\in L(a,b)$ eine in $(a,b)$ fast überall positive Funktion, $\varphi_\lambda(x)\in L^2_w(a,b)$ $(\lambda=1,2,\ldots)$ eine Funktionenfolge, welche kein konstan-*

[11] Insbesondere wird die Gültigkeit einer Differentialgleichung, oder einer anderen speziellen Relation weder bei den Elementen des Systems, noch bei den Gewichtsfunktionen verlangt.

*tes Element enthält.*[12] *Wir nehmen an, daß entweder* 1) $\{\varphi_\lambda(x)\}$ *selbst oder* 2) *die Folge* $\gamma_0, \varphi_1(x), \varphi_2(x), \ldots$, *mit einem konstanten* $\gamma_0 \neq 0$, *ein w-vollständiges und w-orthonormales System*[13] *inbezug auf* $(a, b)$ *bildet.*

I) *Sei* $Q(x)$ *eine in* $(a, b)$ *fast überall positive Funktion,* $Q(x) \in L(a, b)$, *und seien* $c_1, c_2, \ldots$ *gewisse Konstanten; man betrachte die Folge*

$$1,\ \Phi_1(x)+c_1,\ \Phi_2(x)+c_2, \ldots \tag{2.1}$$

*mit*

$$\Phi_\lambda(x) = \int_a^x w(t)\varphi_\lambda(t)\,dt \qquad (\lambda = 1, 2, \ldots). \tag{2.2}$$

*Das Funktionensystem* (2.1) *ist dann und nur dann Q-orthogonal auf* $(a, b)$, *wenn fast überall in* $(a, b)$

$$\varphi_\lambda(x) = \varrho_\lambda \int_a^x Q(t)\,[\Phi_\lambda(t)+c_\lambda]\,dt + K_\lambda \qquad (\lambda = 1, 2, \ldots) \tag{2.3}$$

*gilt, wobei*

$$\varrho_\lambda = -\left\{\int_a^b Q(x)\,[\Phi_\lambda(x)+c_\lambda]^2\,dx\right\}^{-1} \quad \textit{und}$$

$$c_\lambda = -\frac{\int_a^b Q(x)\,\Phi_\lambda(x)\,dx}{\int_a^b Q(x)\,dx} \qquad (\lambda = 1, 2, \ldots) \tag{2.4}$$

*sind, während* $K_\lambda$ *die Null oder* $-\gamma_0^2\varrho_\lambda \int_a^b w(t) \int_a^t Q(u)\,[\Phi_\lambda(u)+c_\lambda]\,du\,dt$ *bedeutet, je nachdem der Fall* 1) *bzw.* 2) *vorliegt.*

II) *Wird* $c_\lambda = 0$ *und* $Q(x)\,\Phi_\lambda(x) \in L(a, b)$ $(\lambda = 1, 2, \ldots)$ *statt* $Q(x) \in L(a, b)$ *vorausgesetzt, so bildet* $\{\Phi_\lambda(x)\}$ *im Falle* 2) *dann und nur dann ein Q-orthogonales System auf* $(a, b)$, *wenn* (2.3) *darin bis auf eine Nullmenge mit* $K_\lambda = -\gamma_0^2\varrho_\lambda \int_a^b w(t) \int_a^t Q(u)\,\Phi_\lambda(u)\,du\,dt$ *und nebst der Bedeutung* (2.4) *von* $\varrho_\lambda$ *gilt.*[14]

[12] Da und später heißen zwei Funktionen nur dann verschieden, wenn ihre Erklärung über einer Menge von positivem Maße abweicht.

[13] D. h. vollständig, usw., in $L_w^2(a, b)$.

[14] Beispiele: 1*—2* alle normierten Sturm—Liouvilleschen Eigenfunktionen erster bzw. zweiter Art, welche zu $[Q(x)^{-1}y']' = \varrho w(x) y$, $Q(x) > 0$, $w(x) > 0$ $(a \leq x \leq b)$ angehören, 3* das trigonometrische Orthonormalsystem (auf einem Intervalle der Länge $2\pi$), 4* die normierten Jacobischen Polynome. Für 1* ist I), Fall 1), für 2*, 3* aber I), Fall 2), während für 4* $(Q(x) = (1-x)^{\alpha+1}(1+x)^{\beta+1}$, $\alpha > -1$, $\beta > -1)$ nur II), Fall 2) anwendbar. (Vgl. [20], S. 265 ff., [29].)

BEWEIS. Es seien alle Voraussetzungen für $w(x)$ und $\{\varphi_\lambda(x)\}$ erfüllt; dann gilt in beiden Fällen 1) und 2)

$$\int_a^b w(x)\varphi_m(x)\varphi_n(x)\,dx = \begin{cases} 1, & \text{falls } m=n, \\ 0 & \text{für } m \neq n \end{cases} \qquad (m,n=1,2,\ldots) \tag{2.5}$$

und im Falle 2) überdies

$$\int_a^b w(x)\varphi_\lambda(x)\,dx = 0 \qquad (\lambda=1,2,\ldots), \tag{2.6}$$

$$\gamma_0^2 \int_a^b w(x)\,dx = 1, \quad \gamma_0^2 = \left[\int_a^b w(x)\,dx\right]^{-1}. \tag{2.7}$$

I) 1° *(Notwendigkeit.)* Nehmen wir an, daß das System (2. 1) unter der Bezeichnung (2. 2) bei gewissen Werten von $c_\lambda$ inbezug auf $(a, b)$ nebst einer angegebenen Belegungsfunktion $Q(x) \in L(a, b)$ orthogonal ist, d. h.

$$\int_a^b Q(x)[\Phi_\lambda(x)+c_\lambda]\,dx = 0 \qquad (\lambda=1,2,\ldots) \tag{2.8}$$

und

$$\int_a^b Q(x)[\Phi_m(x)+c_m][\Phi_n(x)+c_n]\,dx = 0 \qquad (m \neq n;\ m,n=1,2,\ldots). \tag{2.9}$$

Dann können wir aus (2. 8) sofort auf den Wert (2. 4) von $c_\lambda$ $(\lambda=1,2,\ldots)$ schließen; $c_\lambda$ soll also weiterhin zur kurzen Bezeichnung des dortigen Ausdrucks dienen.

Die linke Seite von (2. 9) integrieren wir nach Teilen und zwar in der für Lebesguesche Integrale passenden Weise.[15] Man erhält

$$\begin{aligned} \int_a^b Q(x)[\Phi_m(x)+c_m][\Phi_n(x)+c_n]\,dx &= \int_a^b Q(t)[\Phi_m(t)+c_m]\,dt\cdot[\Phi_n(b)+c_n] - \\ &- \int_a^b w(x)\varphi_n(x)\int_a^x Q(t)[\Phi_m(t)+c_m]\,dt\,dx \qquad (m,n=1,2,\ldots), \end{aligned} \tag{2.10}$$

wobei das Glied in der Mitte vermöge (2. 8) verschwindet. Durch Kombination von (2. 10) und (2. 9) ergibt sich also

$$\begin{aligned} &\int_a^b w(x)\varphi_m(x)\int_a^x Q(t)[\Phi_m(t)+c_m]\,dt\,dx = \\ &= -\int_a^b Q(x)[\Phi_m(x)+c_m]^2\,dx \qquad (m=1,2,\ldots) \end{aligned} \tag{2.11}$$

[15] Vgl. [26], S. 54.

für $m=n$ und

$$\int_a^b w(x)\varphi_n(x)\int_a^x Q(t)[\Phi_m(t)+c_m]\,dt\,dx=0 \qquad (m\neq n;\ m,n=1,2,\ldots) \tag{2.12}$$

sonst.

Aus (2.5) und (2.12) folgt nun

$$\int_a^b w(x)\varphi_n(x)\left\{\varphi_m(x)-\varrho_m\int_a^x Q(t)[\Phi_m(t)+c_m]\,dt\right\}dx=0 \tag{2.13}$$
$$(m\neq n;\ m,n=1,2,\ldots)$$

bei *jeder* Wahl der Zahlenfaktoren $\varrho_m$; indem man zweckmäßig auch für $m=n$

$$\int_a^b w(x)\varphi_m(x)\left\{\varphi_m(x)-\varrho_m\int_a^x Q(t)\,[\Phi_m(t)+c_m]\,dt\right\}dx=0 \tag{2.14}$$

setzt, so führt (2.14) mit Rücksicht auf (2.5) und (2.11) zu

$$\varrho_m=-\frac{\int_a^b w(x)\varphi_m(x)^2\,dx}{\int_a^b Q(x)\,[\Phi_m(x)+c_m]^2dx}= \tag{2.15}$$
$$=-\left\{\int_a^b Q(x)\,[\Phi_m(x)+c_m]^2dx\right\}^{-1} \qquad (m=1,2,\ldots),$$

entsprechend (2.4).

Mit dieser (für die Folgenden festgehaltenen) Bedeutung von $\varrho_m$ schreiben wir zur Abkürzung

$$F_m(x)=\varphi_m(x)-\varrho_m\int_a^x Q(t)\,[\Phi_m(t)+c_m]\,dt \qquad (m=1,2,\ldots);$$

dann geht (2.13) in

$$\int_a^b w(x)\varphi_n(x)F_m(x)\,dx=0 \tag{2.16}$$

über, was aber jetzt für *beliebige positiv-ganze* $m, n$ (auch bei $m=n$) gilt.

Wenn man in (2.16) $m$ fix denkt, während $n$ alle natürlichen Zahlen durchlaufen läßt, so sieht man, daß alle verallgemeinerten Fourier-Koeffizienten von $F_m(x)$ inbezug auf $\{\varphi_\lambda(x)\}$ und die Belegungsfunktion $w(x)$ verschwinden. Daraus folgt im Falle 1), daß

$$F_m(x)\equiv 0 \qquad m=1,2,\ldots),$$

d. h.

$$\varphi_m(x)=\varrho_m\int_a^x Q(t)\,[\Phi_m(t)+c_m]\,dt \qquad (m=1,2,\ldots) \tag{2.17}$$

f. ü. in $(a,b)$.

Im Falle 2) folgt aber (z. B. durch die Parsevalsche Formel bzw. durch die Konvergenz „im Mittel" der Fourierentwicklung nach $\{\varphi_\lambda(x)\}$ von $F_m(x)$[16]), daß

$$F_m(x) \equiv \text{const} = K_m,$$

d. h.

$$\varphi_m(x) = \varrho_m \int_a^x Q(t)\,[\Phi_m(t) + c_m]\,dt + K_m \qquad (m = 1, 2, \ldots) \tag{2.18}$$

f. ü. in $(a, b)$.

(2. 17) und (2. 18) ergeben die erste Hälfte der Behauptung I).

2° *(Hinreichen.)* Umgekehrt, es sei die Gültigkeit von (2. 3) f. ü. in $(a, b)$ und mit der im Satze gegebenen Bedeutung der Konstanten $c_\lambda, \varrho_\lambda, K_\lambda$ angenommen.

Dann folgt (2. 8) aus der Wahl[9] von $c_\lambda$ $(\lambda = 1, 2, \ldots)$, so daß die oben ausgeführte partielle Integration wieder die Formel

$$\int_a^b Q(x)\,[\Phi_m(x) + c_m]\,[\Phi_n(x) + c_n]\,dx =$$

$$= -\int_a^b w(x)\varphi_n(x) \int_a^x Q(t)\,[\Phi_m(t) + c_m]\,dt\,dx \qquad (m, n = 1, 2, \ldots)$$

ergibt. Nun ist aber vermöge (2. 3) und (2. 5) in beiden Fällen

$$\varrho_m \int_a^b w(x)\,\varphi_n(x) \int_a^x Q(t)\,[\Phi_m(t) + c_m]\,dt\,dx = \int_a^b w(x)\varphi_n(x)\,[\varphi_m(x) - K_m]\,dx =$$

$$- K_m \int_a^b w(x)\varphi_n(x)\,dx = 0 \qquad (m \neq n;\ m, n = 1, 2, \ldots), \tag{2.19}$$

weil im vorletzten Gliede bei 1) der erste, bei 2) nach (2. 6) der zweite Faktor verschwindet.

Beachtet man noch, daß offenbar $\varrho_m < 0$ $(m = 1, 2, \ldots)$ ist, so kann man aus den beiden letzten Relationen auf (2. 9), d. h. auf die behauptete Orthogonalitätseigenschaft von (2. 1) schließen.

II) Wird $c_\lambda = 0$ $(\lambda = 1, 2, \ldots)$ gewählt, und statt $Q(x)$ nur $Q(x)\,|\Phi_\lambda(x)|$ $(\lambda = 1, 2, \ldots)$ in $[a, b]$ $L$-integrabel angenommen, ferner der Fall 2) betrachtet, so fällt (2. 8) weg, das mittlere Glied in (2. 10) verschwindet aber wegen (2. 6) wieder; es folgt also, daß alle weiteren Formeln und Feststellungen mit $c_\lambda = 0$ stehen bleiben.

Damit ist Satz 1 bewiesen.

[16] Vgl. z. B. [16], S. 16.

Wir bemerken, daß alle Integrale im Satze 1 sich bequem als diejenigen von Lebesgue—Stieltjes schreiben lassen, was offenbar gewisse Verallgemeinerungen ermöglicht.

Bezeichnen $w^*(x)$ und $Q^*(x)$ unbestimmte Integrale von $w(x)$ bzw. $Q(x)$ in $(a, b)$, so ist

$$\Phi_\lambda(x) = \int_a^x \varphi_\lambda(t)\,dw^*(t) \qquad (\lambda = 1, 2, \ldots), \tag{2.20}$$

und der Teil I) unseres Satzes lautet: *falls* $\int_a^b dQ^*(x)$ *vorhanden und positiv ist, bildet* $1, \{\Phi_\lambda(x) + c_\lambda\}$ *dann und nur dann ein Orthogonalsystem inbezug auf die Verteilung* $dQ^*(x)$ *und* $(a, b)$, *wenn über einer Menge* $E \in (a, b)$, *welche fast alle Punkte* $x$ *von* $(a, b)$ *enthält,*

$$\varphi_\lambda(x) = \varrho_\lambda \int_{x_0}^x [\Phi_\lambda(t) + c_\lambda]\,dt + \varphi_\lambda(x_0) \qquad (\lambda = 1, 2, \ldots) \tag{2.21}$$

*gilt, wobei*

$$\begin{aligned} \varrho_\lambda &= -\left\{\int_a^b [\Phi_\lambda(x) + c_\lambda]^2\,dQ^*(x)\right\}^{-1}, \\ c_\lambda &= \left[\int_a^b dQ^*(x)\right]^{-1} \int_a^b \Phi_\lambda(x)\,dQ^*(x) \qquad (\lambda = 1, 2, \ldots) \end{aligned} \tag{2.22}$$

*und im Falle* 1) $x_0 = a$, $\varphi_\lambda(a) = 0$ $(\lambda = 1, 2, \ldots)$, *während im Falle* 2) $x_0$ *beliebig in* $E$ *gewählt werden soll.*

Teil II) erlaubt die folgende Verschärfung: *ist* $c_\lambda = 0$ $(\lambda = 1, 2, \ldots)$ *und strebt das Produkt* $\Phi_n(x) \int_{x_0}^x \Phi_m(t)\,dQ^*(t)$ *für jedes feste* $m, n, x_0 \in (a, b)$ *für* $x \to a + 0$, $x \to b - 0$ *gegen Null, so bildet* $\{\Phi_\lambda(x)\}$ *im Falle* 2) *genau dann ein Orthogonalsystem, wenn* (2.21) *für jedes* $x \in E$, $x_0 \in E$ *mit* $c_\lambda = 0$, $\varrho_\lambda = -\left[\int_a^b \Phi_\lambda(x)^2\,dQ^*(x)\right]^{-1}$ $(\lambda = 1, 2, \ldots)$ *erfüllt ist.*[17]

**2. 2.** Wenn die Funktionen $\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$, $w(x)$, $Q(x)$ in $(a, b)$ *stetig* und die beiden letzten noch positiv vorausgesetzt werden,[18] so ist dies auf die angewandte Schlußweise nur in dem Sinne von Einfluß, daß in den

[17] Die Bedingungen dieser allgemeineren Fassung II) passen nicht nur für die Jacobischen, sondern auch für die (mit unendlichem $(a, b)$ verknüpften) Laguerreschen bzw. Hermiteschen Polynome.

[18] Da die Stetigkeit in den Endpunkten nicht verlangt wird, sollen die Integrabilitätsforderungen ja für sich beibehalten werden.

obigen Gleichungen die Ausnahmemengen vom Maße Null wegfallen; dann ist (2. 3) wegen der Differenzierbarkeit der unbestimmten Integrale offenbar äquivalent mit einer Differentialgleichung mit Randwertvorschriften.

Man erhält so den

SATZ 2. *Es sei* $w(x) \in L(a, b)$ *eine in* $(a, b)$ *positive und stetige Funktion,* $\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ *eine Folge von nicht konstanten, in* $[a, b]$ *stetigen Funktionen. Wir nehmen an, daß entweder* 1) $\{\varphi_\lambda(x)\}$ *selbst oder* 2) *die Folge* $\gamma_0, \varphi_1(x), \varphi_2(x), \ldots$ *mit einem konstanten* $\gamma_0 \neq 0$, *ein w-vollständiges und w-orthonormales System inbezug auf* $(a, b)$ *bildet.*

I) *Es bezeichne* $Q(x) \in L(a, b)$ *noch eine in* $(a, b)$ *positive und stetige Funktion;* $c_1, c_2, \ldots$ *seien gewisse Konstanten. Man betrachte die Folge:* 1, $\Phi_1(x) + c_1$, $\Phi_2(x) + c_2, \ldots$ *mit* (2. 2).

*Das letztgenannte System ist dann und nur dann Q-orthogonal für* $(a, b)$, *wenn* 1° $\varphi_\lambda(x)$ *und sogar* $Q(x)^{-1}\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ *darin stetig diffenrenzierbar,* 2° *die Relationen*

$$\left[\frac{1}{Q(x)}\varphi_\lambda'(x)\right]' = \varrho_\lambda w(x)\varphi_\lambda(x) \qquad (a < x < b;\ \lambda = 1, 2, \ldots), \tag{2.23}$$

*und* 3° *die Randbedingungen*

$$\left.\begin{aligned} &\varphi_\lambda(a+0) = K_\lambda && (2.24)\\ &\lim_{x \to a+0} \frac{\varphi_\lambda'(x)}{Q(x)} = \varrho_\lambda c_\lambda && (2.25) \end{aligned}\right\} \qquad (\lambda = 1, 2, \ldots)$$

*mit der im Satze* 1 *enthaltenen Bedeutung von* $K_\lambda, \varrho_\lambda, c_\lambda$ *erfüllt sind.*[19]

II) *Ist* $c_\lambda = 0$ $(\lambda = 1, 2, \ldots)$ *und wird die Integrabilität von* $Q(x)|\Phi_\lambda(x)|$ $(\lambda = 1, 2, \ldots)$ *in* $[a, b]$, *statt derjenigen von* $Q(x)$ *angenommen, so bildet nebst* 2) $\{\Phi_\lambda(x)\}$ *genau dann ein Q-orthogonales System für* $(a, b)$, *falls* (2. 23) *und* $\lim\limits_{x \to a+0} [Q(x)^{-1}\varphi_\lambda'(x)] = 0$ $(\lambda = 1, 2, \ldots)$ *gelten.*[20]

**2. 3.** Was die Frage II der Einleitung betrifft, so zeigen wir den

SATZ 3. *Es sei* $w(x) \in L(a, b)$ *eine in* $(a, b)$ *positive und stetige Funktion,* $\varphi_\lambda(x) \in L_w^2(a, b)$ $(\lambda = 1, 2, \ldots)$ *eine Folge ebenda zweimal stetig differenzierbarer Funktionen. Wir nehmen an, daß* $\{\varphi_\lambda(x)\}$ *ein w-vollständiges, w-orthogonales System inbezug auf* $(a, b)$ *bildet.* $g(x)$ *bezeichne eine in* $(a, b)$ *positive, stetig differenzierbare Funktion mit* $w(x)^{-1}[g(x)\varphi_\lambda'(x)]' \in L_w^2(a, b)$ $(\lambda = 1, 2, \ldots)$ *und*

[19] Es sei hervorgehoben, daß diese Bedingungen 1°—3° offenbar noch $\varphi_\lambda(b-0) = K_\lambda$ und im Falle 2) auch $\lim\limits_{x \to b-0} [Q(x)^{-1}\varphi_\lambda'(x)] = \varrho_\lambda c_\lambda$ $(\lambda = 1, 2, \ldots)$ implizieren.

[20] Für Beispiele s. die Fußnote [14].

*es seien die Relationen*

(2.26) $$\lim_{x\to a+0}[g(x)\varphi_m(x)\varphi'_n(x)] = \lim_{x\to b-0}[g(x)\varphi_m(x)\varphi'_n(x)] \qquad (m, n = 1, 2, \ldots)$$

*(mit endlichen Grenzwerten) befriedigt.*[21]

*Damit nun* $\{\varphi'_\lambda(x)\}$ *ein g-orthogonales System inbezug auf* $(a, b)$ *bildet, ist notwendig und hinreichend, daß die Relationen*

(2.27) $$[g(x)\varphi'_\lambda(x)]' = \varrho_\lambda w(x)\varphi_\lambda(x) \qquad (a < x < b;\ \lambda = 1, 2, \ldots)$$

*mit*

(2.28) $$\varrho_\lambda = -\frac{\int_a^b g(x)\varphi'_\lambda(x)^2\,dx}{\int_a^b w(x)\varphi_\lambda(x)^2\,dx} \qquad (\lambda = 1, 2, \ldots)$$

*bestehen.*[22]

Beweis. Es seien unsere Forderungen erfüllt; dann gilt

(2.29) $$\int_a^b w(x)\varphi_m(x)\,\varphi_n(x)\,dx = 0 \qquad (m \neq n;\ m, n = 1, 2, \ldots).$$

1° *(Notwendigkeit.)* Es werde angenommen, daß $\{\varphi'_\lambda(x)\}$ inbezug auf $(a, b)$ $g$-orthogonal ist, d. h.

(2.30) $$\int_a^b g(x)\varphi'_m(x)\varphi'_n(x)\,dx = 0 \qquad (m \neq n;\ m, n = 1, 2, \ldots).$$

Dann liefert eine Teilintegration des linksstehenden Ausdrucks wegen (2.26) die Formel

(2.31) $$\int_a^b g(x)\varphi'_m(x)\varphi'_n(x)\,dx = -\int_a^b [g(x)\varphi'_m(x)]'\,\varphi_n(x)\,dx \qquad (m, n = 1, 2, \ldots),$$

[21] Die erste Annahme: $w(x)^{-1}[g(x)\varphi'_\lambda(x)]' \in L^2_w(a, b)$ gilt z. B. von selbst, wenn die übrigen verlangten Eigenschaften von $w(x), g(x), \{\varphi_\lambda(x)\}$ im ganzen *(abgeschlossenen)* $[a, b]$ bestehen. — Was (2.26) betrifft, so kann sie in mannigfaltiger Weise erfüllt sein, z. B. 1° wenn jedes $\varphi_\lambda(x)$ nebst der Stetigkeit von $g(x)\varphi'_\lambda(x)$ in $[a, b]$ für $x \to a+0$, $x \to b-0$ gegen 0 strebt, oder 2° durch die Beschränktheit von $\varphi_\lambda(x)$ und

$$\lim_{x\to a+0}[g(x)\varphi'_\lambda(x)] = \lim_{x\to b-0}[g(x)\varphi'_\lambda(x)] = 0 \qquad (\lambda = 1, 2, \ldots),$$

ferner 3° nebst im ganzen $[a, b]$ stetigen $g(x), \varphi_\lambda(x), \varphi'_\lambda(x)$ und $g(a) \neq 0$ durch $\varphi_\lambda(a) = M\varphi_\lambda(b)$, $\varphi'_\lambda(a) = M_1\varphi'_\lambda(b)$, $MM_1 = g(a)^{-1}g(b)$ ($M$ und $M_1$ Konstanten), usw. — Verlangt man statt der Differenzierbarkeitsbedingungen die Totalstetigkeit von $\varphi_\lambda(x)$ und $g(x)\varphi'_\lambda(x)$ in $(a, b)$, so kann eine entsprechende Behauptung mit *fast überall* gültiger (2.27) ausgesprochen werden.

[22] Satz 3 ist bei allen betrachteten klassischen Orthogonalsystemen (einschließlich der Laguerreschen und Hermiteschen Polynome) anwendbar.

woraus man insbesondere für $m=n$

$$(2.32)\qquad \int_a^b [g(x)\varphi_m'(x)]'\varphi_m(x)\,dx = -\int_a^b g(x)\varphi_m'(x)^2\,dx,$$

sonst aber durch (2.30)

$$(2.33)\qquad \int_a^b [g(x)\varphi_m'(x)]'\varphi_n(x)\,dx = 0 \qquad (m \neq n;\ m, n = 1, 2, \ldots)$$

erhält.

Auf Grund von (2.29) und (2.33) gilt nun

$$(2.34)\qquad \int_a^b \{[g(x)\varphi_m'(x)]' - \varrho_m w(x)\varphi_m(x)\}\,\varphi_n(x)\,dx = 0 \quad (m \neq n;\ m, n = 1, 2, \ldots)$$

mit jedem reellen $\varrho_m$ $(m = 1, 2, \ldots)$; verlangt man diese Gleichung auch für $m = n$, so ergibt sich infolge (2.32) eben die Wertbestimmung (2.28) von $\varrho_\lambda$.

Es sei diese Bedeutung von $\varrho_m$ in Evidenz gehalten und der Kürze halber

$$(2.35)\qquad G_m(x) = [g(x)\varphi_m'(x)]' - \varrho_m w(x)\varphi_m(x) \qquad (m = 1, 2, \ldots)$$

gesetzt; dann hat man also

$$(2.36)\qquad \int_a^b \frac{G_m(x)}{w(x)}\,\varphi_n(x)w(x)\,dx = 0 \qquad (m, n = 1, 2, \ldots).$$

Wie man sieht, handelt es sich um die Fourier-Koeffizienten bezüglich $\{\varphi_\lambda(x)\}$ und der Belegung $w(x)$ der Funktion $G_m(x)\,w(x)$, welche für jedes feste $m$ wegen $\varphi_m(x) \in L^2_w(a, b)$ und $w(x)^{-1}[g(x)\varphi_m'(x)]' \in L^2_w(a, b)$ gewiß der Klasse $L^2_w(a, b)$ angehört. Da, der Voraussetzung nach, das genannte System $w$-vollständig ist, hat (2.36) die Relation

$$G_m(x) \equiv 0 \qquad (a < x < b;\ m = 1, 2, \ldots),$$

also, in der Tat,

$$(2.37)\qquad [g(x)\varphi_m'(x)] = \varrho_m w(x)\varphi_m(x) \qquad (a < x < b;\ m = 1, 2, \ldots)$$

zur Folge.

2° *(Hinreichen.)* Gehe man, umgekehrt, von der Hypothese aus, daß (2.27) mit (2.28) besteht.

Dann kommen wir, durch Anwendung von (2.27) und (2.33), welche letzte wegen (2.26) ihre Gültigkeit behält, zur Relation

$$\int_a^b g(x)\varphi_m'(x)\varphi_n'(x)\,dx = -\varrho_m \int_a^b w(x)\varphi_m(x)\varphi_n(x)\,dx;$$

daraus geht aber nach (2.29) sofort die fragliche Orthogonalitätseigenschaft von $\{\varphi_\lambda'(x)\}$ hervor.

## § 3. Gewöhnliche $SL$-Systeme

**3.1.** Im Anschluß zu den Sätzen 2 und 3 ist es zweckmäßig eine Benennung einzuführen.

Definition. *Es seien $g(x)$ und $w(x)$ in $a < x < b$ positive, stetig differenzierbare Funktionen, $\varphi_\lambda(x) \in L^2_w(a, b)$ $(\lambda = 1, 2, \ldots)$ sei eine Folge mit nichtkonstanten, linear unabhängigen, ebenda mit stetigen zweiten Ableitungen versehenen Elementen;*[23] *ferner seien $\varrho_\lambda \neq 0$ $(\lambda = 1, 2, \ldots)$ verschiedene reelle Zahlen.*

*Man sagt, die Folge $\{\varphi_\lambda(x)\}$ sei vom gewöhnlichen Sturm—Liouvilleschen (kurz SL-) Typus (oder ein gewöhnliches SL-System) für $(a, b)$ mit den Eigenwerten $\{\varrho_\lambda\}$ zu den Gewichtsfunktionen $(g(x), w(x))$, wenn 1) die Relationen*

$$[g(x)\varphi_\lambda'(x)]' = \varrho_\lambda w(x)\varphi_\lambda(x) \qquad (a < x < b;\ \lambda = 1, 2, \ldots) \tag{3.1}$$

*bestehen, 2) die Funktionen $R_\lambda(x) = g(x)\varphi_\lambda'(x)$ $(\lambda = 1, 2, \ldots)$ in den Punkten $a$ und $b$ von rechts bzw. links stetig,*[24] *3) die Grenzwerte von $\varphi_m(x)R_n(x)$ $(m, n = 1, 2, \ldots)$ für $x \to a + 0$, $x \to b - 0$ vorhanden und die Randbedingungen*

$$\lim_{x \to a+0} [\varphi_m(x)R_n(x)] = \lim_{x \to b-0} [\varphi_m(x)R_n(x)] \qquad (m, n = 1, 2, \ldots) \tag{3.2}$$

*erfüllt sind.*

Es ist nun mühelos einzusehen, daß die eben definierten Funktionenfolgen eine *solche Klasse orthogonaler Systeme bilden, welche fast alle klassischen Systeme umfaßt* und andererseits, daß man von ihnen *durch Differentiation und Integration wieder Orthogonalsysteme* erhält. Es gilt nämlich der folgende (zugleich die Übersicht der bisherigen Ergebnisse erleichternde)

Satz 4. *Es sei $\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ eine Folge vom gewöhnlichen SL-Typus für $(a, b)$ mit den Eigenwerten $\{\varrho_\lambda\}$ zu den Gewichtsfunktionen $(g(x), w(x))$.*

*Dann gelten die Behauptungen:*

*I) $\{\varphi_\lambda(x)\}$, $\{\varphi_\lambda'(x)\}$ und zugleich $\{\Phi_\lambda(x) + c_\lambda\}$ mit*

$$\Phi_\lambda(x) = \int_a^x w(t)\varphi_\lambda(t)\,dt, \quad c_\lambda = \frac{1}{\varrho_\lambda} R_\lambda(a) \qquad (\lambda = 1, 2, \ldots)^{25} \tag{3.3}$$

[23] $g(x)$ und $w(x)$ mögen also in den Endpunkten $x = a$, $x = b$ zufällig verschwinden; sie können, ebenso wie $\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$, dabei auch unendlich werden.

[24] Wird die (zu starke) Forderung gestellt, daß die Stetigkeit von $g(x)$ und $\varphi_\lambda'(x)$ $(\lambda = 1, 2, \ldots)$ im ganzen $[a, b]$ gilt, so ist 2) natürlich überflüssig.

[25] Gehört $w(t)\varphi_\lambda(t)$ nicht zur Klasse $L(a, b)$, so soll $\int_a^x w(t)\varphi_\lambda(t)\,dt = \lim_{A \to a+0} \int_A^x w(t)\varphi_\lambda(t)\,dt$ $(a < A \leq x < b)$ verstanden werden.

*bilden je ein Orthogonalsystem für* $(a, b)$ *und zwar, der Reihe nach, zu den Belegungsfunktionen* $w(x)$, $g(x)$ *und* $Q(x) = g(x)^{-1}$.

*Wenn* $Q(x) \in L(a, b)$ *und* $\varphi_\lambda(x)$ *im ganzen* $[a, b]$ *stetig mit* $\varphi_\lambda(a) = \varphi_\lambda(b)$ $(\lambda = 1, 2, \ldots)$ *ist, so kann* $\{\Phi_\lambda(x) + c_\lambda\}$, *wenn aber nebst* $w(x) \in L(a, b)$ $R_\lambda(a) = R_\lambda(b)$ $(\lambda = 1, 2, \ldots)$ *besteht, so mag* $\{\varphi_\lambda(x)\}$ *durch* 1 *erweitert werden.*

II) *Es ist*

$$\varrho_\lambda = -\left[\int_a^b g(x)\varphi_\lambda'(x)^2\,dx\right]\left[\int_a^b w(x)\varphi_\lambda(x)^2\,dx\right]^{-1} \qquad (\lambda = 1, 2, \ldots)^{26} \tag{3.4}$$

*oder*

$$\varrho_\lambda = -\left[\int_a^b w(x)\varphi_\lambda(x)^2\,dx\right]\left[\int_a^b Q(x)(\Phi_\lambda(x) + c_\lambda)^2\,dx\right]^{-1} \quad (\lambda = 1, 2, \ldots). \tag{3.5}$$

*Man hat ferner, falls jedes* $\varphi_\lambda(x)$ *in* $[a, b]$ *stetig und* $\int_a^b Q(x)\,dx$ *vorhanden ist,*

$$c_\lambda = \left|\int_a^b Q(x)\,dx\right|^{-1}\left|\frac{\varphi_\lambda(b) - \varphi_\lambda(a)}{\varrho_\lambda} - \int_a^b Q(x)\,\Phi_\lambda(x)\,dx\right| \quad (\lambda = 1, 2, \ldots). \tag{3.6}$$

III) *Sind die Funktionen* $\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ *auch in* $[a, b]$ *stetig, so sind die Folgen* $\{\Phi_\lambda(x) + c_\lambda\}$ *und* $\{R_\lambda(x)\}$ *gleichfalls vom gewöhnlichen SL-Typus für* $(a, b)$ *mit den Eigenwerten* $\{\varrho_\lambda\}$, *und zwar zu den Gewichtsfunktionen* $(w(x)^{-1}$, $g(x)^{-1} = Q(x))$.

Beweis. Wir nehmen an, daß unsere Bedingungen erfüllt sind; es sei $a < A < B < b$.

Wenn man in (3.1) $\lambda = m$, $\lambda = n$ setzt und die entstehenden Gleichungen mit $\varphi_n(x)$ bzw. $\varphi_m(x)$ multipliziert, so folgt durch Substraktion

$$\{g(x)[\varphi_m'(x)\varphi_n(x) - \varphi_m(x)\varphi_n'(x)]\}' = (\varrho_m - \varrho_n)w(x)\varphi_m(x)\varphi_n(x); \tag{3.7}$$

integriert man noch von $A$ bis $B$, so ergibt sich die Relation

$$(\varrho_m - \varrho_n)\int_A^B w(x)\varphi_m(x)\varphi_n(x)\,dx = \varphi_n(x)R_m(x) - \varphi_m(x)R_n(x)\Big|_{x=A}^{x=B} \tag{3.8}$$
$$(m, n = 1, 2, \ldots).$$

Die rechte Seite strebt laut (3.2) für $A \to a + 0$, $B \to b - 0$ gegen Null; somit hat auch das linksstehende Integral einen Limes und wegen $\varrho_m \neq \varrho_n$

[26] Wie aus (3.4) und (3.5) hervorgeht, müssen Zähler und Nenner positiv, d. h. $\varrho_\lambda < 0$ $(\lambda = 1, 2, \ldots)$ sein. — Bei konkreten Folgen $\{\varphi_\lambda(x)\}$ können somit die Eigenwerte sofort berechnet und z. B. ihr asymptotisches Verhalten untersucht werden. Umgekehrt: sucht man $y = \varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ als Eigenfunktionen des durch $[g(x)y']' = \text{const.}\, y$ $(a < x < b)$ und Randbedingungen vom (3.2)-Typ angegebenen Randwertproblems, so ist das einschlägige *Variationsprinzip* eben auf (3.4) oder (3.5) gegründet. (Vgl. z. B. [24]; [19], S. 214 ff.)

$(m \neq n)$ kann man auf

$$\int_a^b w(x)\varphi_m(x)\varphi_n(x)\,dx = 0 \qquad (m \neq n;\ m, n = 1, 2, \ldots) \tag{3.9}$$

schließen.

Andererseits gibt eine partielle Integration und die Benutzung von (3. 1)

$$\int_A^B g(x)\varphi_m'(x)\varphi_n'(x)\,dx = R_m(x)\varphi_n(x)\Big|_{x=A}^{x=B} - \varrho_m \int_A^B w(x)\varphi_m(x)\varphi_n(x)\,dx \tag{3.10}$$
$$(m, n = 1, 2, \ldots),$$

woher durch (3. 2) und (3. 9) bei $A \to a+0$, $B \to b-0$

$$\int_a^b g(x)\varphi_m'(x)\varphi_n'(x)\,dx = 0 \qquad (m \neq n;\ m, n = 1, 2, \ldots) \tag{3.11}$$

folgt.

Endlich besteht, wie (3. 1) und (3. 3) zeigt,

$$\varrho_\lambda[\Phi_\lambda(x) + c_\lambda] = \lim_{A \to a+0}\left[\varrho_\lambda \int_A^x w(t)\varphi_\lambda(t)\,dt + R_\lambda(A)\right] = R_\lambda(x) \tag{3.12}$$
$$(a < x < b;\ \lambda = 1, 2, \ldots),$$

so daß man durch (3. 11) und (3. 12) zu

$$\int_a^b Q(x)[\Phi_m(x) + c_m][\Phi_n(x) + c_n]\,dx = \frac{1}{\varrho_m \varrho_n}\int_a^b g(x)\varphi_m'(x)\varphi_n'(x)\,dx = 0 \tag{3.13}$$
$$(m \neq n;\ m, n = 1, 2, \ldots)$$

gelangt. — (3. 9), (3. 11), (3. 13) enthalten eben die fraglichen Orthogonalitätseigenschaften von $\{\varphi_\lambda(x)\}$, $\{\varphi_\lambda'(x)\}$ bzw. $\{\Phi_\lambda(x) + c_\lambda\}$.

(3. 12) impliziert auch die Gleichungen

$$\left.\begin{aligned} &\varrho_\lambda \int_A^B w(t)\varphi_\lambda(t)\,dt = R_\lambda(B) - R_\lambda(A) && (3.14)\\ &\varrho_\lambda \int_A^B Q(t)[\Phi_\lambda(t) + c_\lambda]\,dt = \varphi_\lambda(B) - \varphi_\lambda(A) && (3.15) \end{aligned}\right\} \quad (\lambda = 1, 2, \ldots).$$

Wenn nun $R_\lambda(a) = R_\lambda(b)$ oder $\varphi_\lambda(a) = \varphi_\lambda(b)$ $(\lambda = 1, 2, \ldots)$ gilt, so strebt die erste bzw. zweite Differenz rechter Hand für $A \to a+0$, $B \to b-0$ gegen Null, d. h. das entsprechende Integral über $(a, b)$ muß verschwinden; dies bedeutet aber nichts anderes, als daß das Element 1 nebst $w(x) \in L(a, b)$ bzw. $Q(x) \in L(a, b)$ dem Orthogonalsysteme $\{\varphi_\lambda(x)\}$ bzw. $\{\Phi_\lambda(x) + c_\lambda\}$ angehört.

II) Aus (3. 10) entspringt bei $m=n=\lambda$ nach einem Grenzübergange durch (3. 2)

$$\varrho_\lambda \int_a^b w(x)\varphi_\lambda(x)^2\,dx = -\int_a^b g(x)\varphi_\lambda'(x)^2\,dx \qquad (\lambda=1,2,\ldots), \tag{3. 16}$$

was sofort (3. 4) liefert.[27]

Um (3. 5) zu bekommen, konstatieren wir vor allem auf Grund von (3. 12) die Gleichheit

$$[\Phi_\lambda(x)+c_\lambda]\cdot\left\{\varrho_\lambda \int_A^x Q(t)[\Phi_\lambda(t)+c_\lambda]\,dt + \varphi_\lambda(A)\right\}\Bigg|_{x=A}^{x=B} = \frac{1}{\varrho_\lambda}\,\varphi_\lambda(x)R_\lambda(x)\Bigg|_{x=A}^{x=B} \tag{3. 17}$$
$$(\lambda=1,2,\ldots).$$

Somit gelangt man aber durch Anwendung einer Teilintegration und des Grenzübergangs $A\to a+0$, $B\to b-0$ zur Formel

$$\int_a^b w(x)\varphi_\lambda(x)^2\,dx = -\varrho_\lambda \int_a^b Q(x)[\Phi_\lambda(x)+c_\lambda]^2\,dx \qquad (\lambda=1,2,\ldots), \tag{3. 18}$$

woher der fragliche zweite Ausdruck für $\varrho_\lambda$ ablesbar ist.

Schließlich findet man, von der schon benutzten Relation (vgl. (3. 15))

$$\varrho_\lambda \int_a^b Q(t)[\Phi_\lambda(t)+c_\lambda]\,dt = \varphi_\lambda(b)-\varphi_\lambda(a) \tag{3. 19}$$

ausgehend, unmittelbar die Wertbestimmung (3. 6) für $c_\lambda$ ($\lambda=1,2,\ldots$), wofern nur die dort angegebenen Voraussetzungen gemacht werden.

III) Seien $\varphi_\lambda(x)$ ($\lambda=1,2,\ldots$) in $[a,b]$ stetig angenommen; die Integrale

$$\int_a^b Q(x)R_\lambda(x)^2\,dx = \varrho_\lambda^2 \int_a^b Q(x)[\Phi_\lambda(x)+c_\lambda]^2\,dx$$

sind gleich $\int_a^b g(x)\varphi_\lambda'(x)^2\,dx$ ($\lambda=1,2,\ldots$), existieren also laut (3. 16) nach Hypothese.

In diesem Falle ergeben (3. 1) und (3. 2) durch eine einfache Rechnung, daß auch die Folgen $\{R_\lambda(x)\}$ und $\{\Phi_\lambda(x)+c_\lambda\}$ die verlangten Eigenschaften 1)—3) gewöhnlicher $SL$-Systeme besitzen, w. z. b. w.

**3. 2.** Der Teil III) des eben bewiesenen Satzes legt die Frage nahe: *wann hat ein gewöhnliches $SL$-System eine Ableitungsfolge bzw. höhere Deriviertenfolge derselben Art?*[28]

[27] Das linksstehende Integral existiert nach Hypothese!

[28] Wie man weiß, liefern $\{\sin\lambda x\}$, $\{\cos\lambda x\}$, ferner die Jacobischen, Hermiteschen, Laguerreschen Polynome je Beispiele solcher $SL$-Systeme.

In dieser Richtung stützt man sich vor allem auf die folgende Bemerkung, welche sich durch einfache Umformungen verifizieren läßt:

Lemma 1. *Genügt eine in $(a, b)$ dreimal stetig differenzierbare Funktion $y=\varphi_\lambda(x)$ nebst ebenda einmal bzw. zweimal stetig differenzierbarer, positiver $w(x)$ und $g(x)$ der Gleichung (3. 1), so ist $y=\varphi'_\lambda(x)$ eine Lösung der Sturm—Liouvilleschen Differentialgleichung*

$$\left[\frac{g(x)^2}{w(x)}y'\right]'=\left\{\varrho_\lambda-\left[\frac{g'(x)}{w(x)}\right]'\right\}g(x)y \qquad (a<x<b). \tag{3. 20}$$

(3. 20) setzt einerseits in Evidenz die $g$-Orthogonalität von $\{\varphi'_\lambda(x)\}$, falls $\{\varphi_\lambda(x)\}$ eine gewöhnliche $SL$-Folge zu $(g(x), w(x))$ bildet (vgl. Satz 4, I)), andererseits aber zeigt sie, daß eine besondere Wahl der sukzessiven Gewichtsfunktionen, nämlich der Fall, wo diese (wie $w(x)$, $g(x)$, $g(x)^2w(x)^{-1}$) *eine geometrische Progression* bilden, bei $SL$-Systemen ausgezeichnet ist.

In der Tat, gilt

Satz 5. *Es sei $\{\varphi_\lambda(x)\}$ vom gewöhnlichen $SL$-Typus für $(a, b)$ mit den Eigenwerten $\{\varrho_\lambda\}$ zu den Gewichtsfunktionen $(g(x), w(x))$; $g_k(x)$ $(k=1, 2, \ldots)$ seien in $(a, b)$ positive, zweimal stetig differenzierbare Funktionen derart, daß $w(x), g_1(x), g_2(x), \ldots$ eine geometrische Progression bildet. Wir nehmen an, daß die zur Anwendung kommenden Ableitungen von $\varphi_\lambda(x)$ $(\lambda=1, 2, \ldots)$ in $(a, b)$ vorhanden und stetig sind, und ferner, daß die Folgen $\{\varphi_\lambda^{(k)}(x)\}$, für jedes in Betracht kommende $k$, keine Konstante enthalten.*

I) *Dann hat man dafür, damit auch $\{\varphi'_\lambda(x)\}$ eine gewöhnliche $SL$-Folge für $(a, b)$ zu $(g_2(x), g_1(x))$ bildet, folgende notwendige und hinreichende Bedingungen:*

1) *Es sei $g'_1(x)w(x)^{-1}=Px+Q$ $(a<x<b)$ mit konstanten $P$ und $Q$; 2) die Funktionen $g_2(x)\varphi''_\lambda(x)$ $(\lambda=1, 2, \ldots)$ seien für $x\to a+0$, $x\to b-0$ stetig; 3) das Produkt $(Px+Q)g_1(x)\varphi'_\lambda(x)^2$ habe denselben Limes $\eta_\lambda$ $(\lambda=1, 2, \ldots)$ für $x\to a+0$, $x\to b-0$; bei jedem $\lambda$, wofür $\eta_\lambda\neq 0$ ist, besitze $\lim\limits_{x\to a+0}\operatorname{sg}\varphi'_\lambda(x)\cdot\lim\limits_{x\to b-0}\operatorname{sg}\varphi'_\lambda(x)$ denselben Wert.*[29]

*Unter diesen Voraussetzungen ist die zu $\{\varphi'_\lambda(x)\}$ gehörige Eigenwertfolge $\{\varrho_\lambda-P\}$.*

II) *Man wähle $\nu\geqq 2$. Die Systeme $\{\varphi'_\lambda(x)\}, \{\varphi''_\lambda(x)\}, \ldots, \{\varphi_\lambda^{(\nu)}(x)\}$ sind dann und nur dann gleichzeitig vom gewöhnlichen $SL$-Typus und zwar, der Reihe nach, zu den Gewichtsfunktionen $(g_2(x), g_1(x)), (g_3(x), g_2(x)), \ldots, (g_{\nu+1}(x), g_\nu(x))$, wenn die folgenden Bedingungen erfüllt sind:*

[29] Dies bedeutet offenbar, daß jedes $\varphi'_\lambda(x)$ (für die betreffenden $\lambda$) in der Nähe von $a$ und $b$ das gleiche Vorzeichen hat.

1) $g_1'(x)w(x)^{-1} = Px + Q,\quad g_1(x)w(x)^{-1} = px^2 + qx + r\quad (a < x < b)$ *mit konstanten* $P, Q, p, q, r$; 2) $g_1(x)\left(\frac{g_1(x)}{w(x)}\right)^k \varphi_\lambda^{(k+1)}(x)$ *ist in den Punkten a und b von rechts bzw. links stetig* $(k = 1, 2, \ldots, \nu;\ \lambda = 1, 2, \ldots)$; 3) *der Ausdruck*

$$g_1(x)\left(\frac{g_1(x)}{w(x)}\right)^{k-1}\left[\frac{g_1'(x)}{w(x)} + (k-1)\left(\frac{g_1(x)}{w(x)}\right)'\right]\varphi_\lambda^{(k)}(x)^2$$

*strebt gegen denselben (endlichen) Grenzwert* $\eta_{\lambda k}$ $(k = 1, 2, \ldots, \nu;\ \lambda = 1, 2, \ldots)$ *für* $x \to a + 0,\ x \to b - 0$; 4) *bei jedem fixen k und sämtlichen* $\lambda$ *mit* $\eta_{\lambda k} \neq 0$ *ist das Produkt* $\lim\limits_{x\to a+0} \operatorname{sg} \varphi_\lambda^{(k)}(x) \cdot \lim\limits_{x\to b-0} \operatorname{sg} \varphi_\lambda^{(k)}(x)$ *konstant* $(k = 1, 2, \ldots, \nu)$.

*Unter diesen Annahmen ist die mit dem Systeme* $\{\varphi_\lambda^{(k)}(x)\}$ *verknüpfte Eigenwertfolge* $\{\varrho_\lambda - kP - k(k-1)p\}$ $(k = 1, 2, \ldots, \nu)$.

Korollarien. (*) Die Anwendung des Satzes 4 ergibt u. a.: neben $\{\varphi_\lambda(x)\}$ und $\{\varphi_\lambda'(x)\}$ ist auch $\{\varphi_\lambda''(x)\}$ orthogonal für $(a, b)$ (zur Belegung $g_2(x)$), falls die Bedingungen II) 1)—4) erfüllt sind, ferner bilden ja auch die höheren Ableitungen bis zu $\varphi_\lambda^{(\nu)}(x)$ $(\lambda = 1, 2, \ldots)$ Orthogonalsysteme für $(a, b)$ (zu den Belegungen $g_k(x)$; $k = 2, 3, \ldots, \nu$), wenn die Forderungen III) 1)—4)) befriedigt sind.

(**) Folgende Bedingungen sind offenbar *hinreichend*, weil sie die Erfüllung der entsprechenden oben erwähnten Bedingungen nach sich ziehen: 1) $g_1(x)$ sei eine Konstante und 2) $\varphi_\lambda''(x)w(x)^{-1}$ $(\lambda = 1, 2, \ldots)$ für $x \to a + 0, x \to b - 0$ stetig; oder samt II) 1) und 2) noch drittens: $\lim\,[g_1(x)\varphi_\lambda'(x)^2]$ und $\lim\,[xg_1(x)\varphi_\lambda'(x)^2]$ für $x \to a + 0,\ x \to b - 0$ verschwinde (beide statt II) 1)—4); neben III) 1) und 2) noch drittens: $x^m g_1(x)\varphi_\lambda^{(k)}(x)^2$ habe den Grenzwert Null für $x \to a + 0,\ x \to b - 0$ bei allen Werten $m = 0, 1, \ldots, 2\nu - 1$; $k = 1, 2, \ldots, \nu$ (statt III) 1)—4)).

(***) Man erkennt leicht, daß die Bedingung III) 1) folgendes impliziert: $g_1(x)$ soll eine Konstante oder ein Ausdruck der Form

$$\exp\left(\int \frac{Px + Q}{px^2 + qx + r}\,dx\right)$$

sein, d. h. sich jedenfalls mit Hilfe von rationalen Funktionen, Logarithmus bzw. Arcustangens und Exponentialfunktion aufbauen lassen, während andererteils die Darstellung $w(x) = (px^2 + qx + r)^{-1}g_1(x)$ gelten muß.

Beweis. Es seien die über die $SL$-Folge[30] $\{\varphi_\lambda(x)\}$, ferner über $\{\varrho_\lambda\}$, $g_1(x), w(x)$ gemachten Forderungen in Evidenz gehalten.

I) 1° *(Hinreichen.)* Wir nehmen an, daß auch die Bedingungen I) 1)—4) befriedigt sind und haben zu zeigen, daß $\{\varphi_\lambda'(x)\}$ die Eigenschaften einer

[30] Wir benutzen diese Bezeichnung, der Kürze halber, statt „Folge vom gewöhnlichen $SL$-Typus".

$SL$-Folge für $(a, b)$ zu $(g_2(x), g_1(x))$ besitzt. (Vgl. die oben angegebene Definition.)

Vor allem bestätigt man, daß (3. 16) nach Hypothese besteht, womit die Existenz der Integrale $\int_a^b g(x)\varphi'_\lambda(x)^2 dx$ $(\lambda = 1, 2, \ldots)$ gesichert ist.

Wegen I) 1) ergibt nun (3. 20)

$$[g_2(x)\varphi''_\lambda(x)]' = (\varrho_\lambda - P)g_1(x)\varphi'_\lambda(x), \tag{3.21}$$

entsprechend (3. 1);[31] andererteils hat man die Identität für jedes $m, n$

$$g_2(x)\varphi'_m(x)\varphi''_n(x) = \frac{g_1(x)}{w(x)}\varphi'_m(x)\{[g_1(x)\varphi'_n(x)]' - g'_1(x)\varphi'_n(x)\}, \tag{3.22}$$

welche durch (3. 1) und die Bedingung I) 1) in

$$g_2(x)\varphi'_m(x)\varphi''_n(x) = \varrho_n g_1(x)\varphi_n(x)\varphi'_m(x) - g_1(x)(Px+Q)\varphi'_m(x)\varphi'_n(x) \tag{3.23}$$

transformiert werden kann. Das erste Glied rechter Hand hat nach (3. 2) übereinstimmende (endliche) Grenzwerte für $x \to a+0$ und $x \to b-0$; dasselbe gilt bezüglich des zweiten Gliedes. Denn zunächst ist auf Grund von I) 3) (angewendet mit $\lambda = m, \lambda = n$)

$$\begin{aligned} &\lim_{x\to a+0} [g_1(x)^2(Px+Q)^2\varphi'_m(x)^2\varphi'_n(x)^2] = \\ &= \lim_{x\to b-0} [g_1(x)^2(Px+Q)^2\varphi'_m(x)^2\varphi'_n(x)^2] = \eta_m\eta_n, \end{aligned} \tag{3.24}$$

also jedenfalls

$$\begin{aligned} &\lim_{x\to a+0} |g_1(x)(Px+Q)\varphi'_m(x)\varphi'_n(x)| = \\ &= \lim_{x\to b-0} |g_1(x)(Px+Q)\varphi'_m(x)\varphi'_n(x)| = \sqrt{\eta_m\eta_n}. \end{aligned} \tag{3.25}$$

Ist $\eta_m\eta_n = 0$, so kommt man unmittelbar zu

$$\lim_{x\to a+0} [g_1(x)(Px+Q)\varphi'_m(x)\varphi'_n(x)] = \lim_{x\to b-0} [g_1(x)(Px+Q)\varphi'_m(x)\varphi'_n(x)]. \tag{3.26}$$

Widrigenfalls (d. h. bei $\eta_m\eta_n > 0$) kann weder $\operatorname{sg}\varphi'_m(x)$, noch $\operatorname{sg}\varphi'_n(x)$ in den Endpunkten $a$ und $b$ den Grenzwert Null haben, so daß aus I) 4)

$$\frac{\lim\limits_{x\to a+0} \operatorname{sg}\varphi'_m(x)}{\lim\limits_{x\to b-0} \operatorname{sg}\varphi'_m(x)} = \frac{\lim\limits_{x\to a+0} \operatorname{sg}\varphi'_n(x)}{\lim\limits_{x\to b-0} \operatorname{sg}\varphi'_n(x)} \tag{3.27}$$

folgt und sogar, da wegen I) 3)

$$\lim_{x\to a+0} \operatorname{sg}(Px+Q) = \lim_{x\to b-0} \operatorname{sg}(Px+Q) \tag{3.28}$$

[31] Dies zeigt zugleich, die zu $\{\varphi'_\lambda(x)\}$ gehörige Eigenwertfolge sei $\{\varrho_\lambda - P\}$, wie behauptet.

besteht, auch

$$\lim_{x\to a+0} \operatorname{sg}(g_1(x)(Px+Q)\varphi'_m(x)\varphi'_n(x)) = \tag{3.29}$$
$$= \lim_{x\to b-0} \operatorname{sg}(g_1(x)(Px+Q)\varphi'_m(x)\varphi'_n(x)).$$

Aus (3.29) und (3.25) folgt wieder (3.26).

Somit gelangt man endlich durch (3.23) zu

$$\lim_{x\to a+0}[g_2(x)\varphi'_m(x)\varphi''_n(x)] = \lim_{x\to b-0}[g_2(x)\varphi'_m(x)\varphi''_n(x)] \tag{3.30}$$
$$(a<x<b;\ m, n=1, 2, \ldots).$$

Unter Hinzunahme der Bedingung I) 2) sieht man also, daß sämtliche fragliche Eigenschaften von $\{\varphi'_\lambda(x)\}$ bestehen.

2° *(Notwendigkeit.)* Sei $\{\varphi'_\lambda(x)\}$ als ein System vom gewöhnlichen *SL*-Typus für $(a, b)$ zu $(g_2(x), g_1(x))$ angenommen.

Dann muß für jedes $\lambda$ eine Relation der Form

$$[g_2(x)\varphi''_\lambda(x)]' = C_\lambda g_1(x)\varphi'_\lambda(x) \qquad (a<x<b;\ C_\lambda=\text{const.}), \tag{3.31}$$

ferner (3.30), sowie die einseitige Stetigkeit der Funktionen $g_2(x)\varphi''_\lambda(x)$ $(\lambda=1, 2, \ldots)$ in $a$ und $b$ gelten.

Das letzte ist eben die Bedingung I) 2); der Vergleich von (3.20) und (3.31) ergibt die Forderung: $\left[\dfrac{g'_1(x)}{w(x)}\right]' = \text{const.}$, d. h. I) 1). Um noch I) 3) und 4) zu erhalten, fassen wir die aus (3.22) und (3.1) fließende Formel:

$$g_2(x)\varphi'_m(x)\varphi''_n(x) = \varrho_n g_1(x)\varphi_n(x)\varphi'_m(x) - g_1(x)\frac{g'_1(x)}{w(x)}\varphi'_m(x)\varphi'_n(x) \tag{3.32}$$
$$(a<x<b;\ m, n=1, 2, \ldots)$$

ins Auge. Daraus folgt für $m=n=\lambda$ $(\lambda=1, 2, \ldots)$

$$g_1(x)\frac{g'_1(x)}{w(x)}\varphi'_\lambda(x)^2 = \varrho_\lambda g_1(x)\varphi_\lambda(x)\varphi'_\lambda(x) - g_2(x)\varphi'_\lambda(x)\varphi''_\lambda(x), \tag{3.33}$$

wobei (nach Hypothese) sowohl das zweite, als das dritte Glied für $x\to a+0$, $x\to b-0$ übereinstimmende (endliche) Grenzwerte hat. Somit kann man unter Beachtung von I) 1) auf I) 3) schließen.

Gleichfalls impliziert aber (3.32) die Gleichungen

$$\lim_{x\to a+0}[g_1(x)(Px+Q)\varphi'_m(x)\varphi'_n(x)] = \tag{3.34}$$
$$= \lim_{x\to b-0}[g_1(x)(Px+Q)\varphi'_m(x)\varphi'_n(x)] = \xi_{mn} \qquad (m, n=1, 2, \ldots).$$

Sind nun $m$ und $n$ zwei beliebige natürliche Zahlen mit $\eta_m \neq 0$, $\eta_n \neq 0$, so hat man durch (3.24) $\xi_{mn} \neq 0$ und wegen (3.34)

$$\lim_{x\to a+0} \operatorname{sg}(g_1(x)(Px+Q)\varphi'_m(x)\varphi'_n(x)) = \tag{3.35}$$
$$= \lim_{x\to b-0} \operatorname{sg}(g_1(x)(Px+Q)\varphi'_m(x)\varphi'_n(x)) \neq 0;$$

da dies für $m = n$ zeigt, daß

$$\lim_{x\to a+0} \mathrm{sg}\,(g_1(x)(Px+Q)) = \lim_{x\to b-0} \mathrm{sg}\,(g_1(x)(Px+Q))$$

sein soll, erhalten wir

$$\frac{\lim\limits_{x\to a+0} \mathrm{sg}\,\varphi_m'(x)}{\lim\limits_{x\to b-0} \mathrm{sg}\,\varphi_m'(x)} = \frac{\lim\limits_{x\to b-0} \mathrm{sg}\,\varphi_n'(x)}{\lim\limits_{x\to a+0} \mathrm{sg}\,\varphi_n'(x)}$$

oder

$$\lim_{x\to a+0} \mathrm{sg}\,\varphi_m'(x)\cdot \lim_{x\to b-0} \mathrm{sg}\,\varphi_m'(x) = \lim_{x\to a+0} \mathrm{sg}\,\varphi_n'(x)\cdot \lim_{x\to b-0} \mathrm{sg}\,\varphi_n'(x). \tag{3.36}$$

(3. 36) besagt dasselbe, als I) 4).

II) Die Untersuchung des zweiten Teils begründen wir zweckmäßig auf den folgenden kleinen *Hilfssatz:*

Es sei $k$ positiv-ganz und $\{\varphi_\lambda^{(k-1)}(x)\}$ eine Folge vom gewöhnlichen $SL$-Typus für $(a, b)$ zu $(g_k(x), g_{k-1}(x))$. In diesem Falle bildet $\{\varphi_\lambda^{(k)}(x)\}$ dann und nur dann eine $SL$-Folge für $(a, b)$ zu $(g_{k+1}(x), g_k(x))$, wenn $g_1'(x)w(x)^{-1} + (k-1)[g_1(x)w(x)^{-1}]'$ ein Polynom höchstens ersten Grades ist und die Bedingungen II) 2)—4) *für das betreffende* $k$ bestehen.

In der Tat, wendet man den soeben bewiesenen Teil I) auf das System $\{\varphi_\lambda^{(k-1)}(x)\}$ an, wobei $g_k(x)$, $g_{k-1}(x)$ statt der dortigen $g_1(x)$ bzw. $w(x)$ geschrieben werden sollen, so ergibt sich unmittelbar die Notwendigkeit und Hinlänglichkeit der genannten vier Bedingungen.

Nun ist der Fall $\nu = 2$ der Behauptung II) leicht zu erledigen.

1° Sei angenommen, daß II) 1)—4) für $k = 1$, $k = 2$ erfüllt sind. Aus II) 1) und II) 2)—4) mit $k = 1$ können wir gemäß des Teils I) schließen, $\{\varphi_\lambda'(x)\}$ sei eine $SL$-Folge für $(a, b)$ zu $(g_2(x), g_1(x))$; dies und die Bedingungen II) 1), II) 2)—4) mit $k = 2$ involvieren aber nach dem Hilfssatze, daß auch $\{\varphi_\lambda''(x)\}$ ein gewöhnliches $SL$-System für $(a, b)$ zu $(g_3(x), g_2(x))$ bildet.

2° Setzen wir voraus, $\{\varphi_\lambda'(x)\}$ und $\{\varphi_\lambda''(x)\}$ seien vom gewöhnlichen $SL$-Typus für $(a, b)$ zu $(g_2(x), g_1(x))$ bzw. $(g_3(x), g_2(x))$. Dann ergibt der Teil I) die Gültigkeit der Bedingungen I) 1)—4), welche durch den Hilfssatz zu II) 2)—4) mit $k = 1$, $k = 2$, ferner zu

$$\frac{g_1'(x)}{w(x)} = \text{linear}, \qquad \frac{g_1'(x)}{w(x)} + \left[\frac{g_1'(x)}{w(x)}\right]' = \text{linear} \tag{3.37}$$

erweitert werden können. Die beiden letzten liefern II) 1).

Ist $\nu > 2$, so führt eine Induktion nebst Benutzung des Hilfssatzes einfach zum Ziele.

Endlich ist es nicht schwer, unsere Behauptung über die Eigenwerte zu verifizieren.

Es sei $\{\varphi_\lambda^{(k-1)}(x)\}$ eine $SL$-Folge für $(a,b)$ zu $(g_k(x), g_{k-1}(x))$ und $\{\varphi_\lambda^{(k)}(x)\}$ auch eine solche zu $(g_{k+1}(x), g_k(x))$ $(k=1,2,\dots,\nu)$; dann sieht man durch Anwendung des Teils II) sofort ein, daß die zu $\{\varphi_\lambda^{(k)}(x)\}$ gehörige Eigenwertfolge nur $\{\varrho_\lambda - kP - k(k-1)p\}$ sein kann. $y=\varphi_\lambda^{(k)}(x)$ genügt also der Differentialgleichung

$$(3.38)\quad [g_{k+1}(x)\varphi_\lambda^{(k+1)}(x)]' = [\varrho_\lambda - kP - k(k-1)p]g_k(x)\varphi_\lambda^{(k)}(x) \quad (k=1,2,\dots,\nu).$$

## § 4. Anwendungen: Charakterisierung der „klassischen" Orthogonalsysteme

**4. 1.** Unsere bisherigen Sätze zeigen, daß ein Orthogonalsystem $\{\varphi_\lambda(x)\}$ stetiger Funktionen, welches die betreffenden Eigenschaften besitzt, nur aus solchen Elementen bestehen kann, die einer Sturm—Liouvilleschen Differentialgleichung genügen; andererteils, wie erwähnt, haben jedenfalls die klassischen orthogonalen Systeme einer oder andere jener (mit Orthogonalität, Differentiation und Integration verknüpften) Eigenschaften.

Dies ermöglicht uns offenbar, eine Klasse von Sturm—Liouvilleschen Eigenfunktionen, insbesondere das trigonometrische System und die klassischen orthogonalen Polynome in einfacher Weise zu charakterisieren.

SATZ 6. *Es seien $g(x), w(x)$ in einem endlichen Intervalle $(a,b)$ stetig differenzierbare und positive*[32] *Funktionen mit $g(x)^{-1}\in L(a,b)$, $w(x)\in L(a,b)$, für welche eines der beiden Randwertprobleme: „gesucht die in $[a,b]$ beschränkten Integrale $y=y(x)$ von*

$$(4.1)\qquad [g(x)y']' = \varrho w(x)y \qquad (a<x<b;\ \varrho \text{ ein reelles Parameter})$$

*mit hierin stetiger $g(x)y'$ und mit* (I) $y(a+0)=y(b-0)=0$ *oder* (II) $\lim\limits_{x\to a+0}[g(x)y'(x)] = \lim\limits_{x\to b-0}[g(x)y'(x)]=0$" *nicht-triviale Lösungen besitzt. Wir bezeichnen die von Null verschiedenen (einfachen) Eigenwerte des betreffenden Problems mit $(0>)\varrho_1>\varrho_2>\cdots$ und die entsprechenden $w$-normierten Eigenfunktionen, der Reihe nach, mit $\varphi_1(x), \varphi_2(x), \dots$; das letzte System sei bei* (I) *selbst, bei* (II) *samt* $\gamma_0 = \left[\int_a^b w(x)dx\right]^{-\frac{1}{2}} (>0)$ *für $(a,b)$ $w$-abgeschlossen. Man schreibt noch* $\Phi_\lambda(x) = \int_a^x w(t)\varphi_\lambda(t)dt$ $(\lambda=1,2,\dots)$.

*Dann läßt sich $\{\varphi_\lambda(x)\}$ bis auf Vorzeichen und Reihenfolge folgendermaßen charakterisieren: im Falle der Randbedingung* (I) *durch*

1)* $\{\varphi_\lambda(x)\}$ *bildet auf $(a,b)$ ein $w$-vollständiges, $w$-orthonormales System von nicht-konstanten, stetigen Funktionen;*

[32] Das Verschwinden der Gewichtsfunktionen $g(x), w(x)$ in $a$ und $b$ ist also erlaubt.

2)* *das System* $1, \{\Phi_\lambda(x)+c_\lambda\}$ *ist* $g(x)^{-1}$*-orthogonal inbezug auf* $(a, b)$, *wobei* $c_\lambda$ $(\lambda=1, 2, \ldots)$ *geeignet gewählte Konstanten bedeuten;*

*im Falle von* (II) *aber durch*

1)** $\varphi_\lambda(x)$ $(\lambda=1, 2, \ldots)$ *sind in* $(a, b)$ *stetig;* $\gamma_0, \{\varphi_\lambda(x)\}$ *bildet hier ein* $w$*-vollständiges,* $w$*-orthonormales System;*

2)** $\{\Phi_\lambda(x)\}$ *ist* $g(x)^{-1}$*-orthogonal inbezug auf* $(a, b)$.

Beweis. 1° Wir nehmen an, daß alle Bedingungen des Satzes erfüllt sind.

Dann ist $\{\varphi_\lambda(x)\}$ *ein gewöhnliches SL-System,* verknüpft mit der Eigenwertfolge $\{\varrho_\lambda\}$ und den Belegungsfunktionen $(g(x), w(x))$, so daß Satz 4 angewendet werden kann. Unter Beachtung der Annahmen bezüglich der Normiertheit und Abgeschlossenheit von $\{\varphi_\lambda(x)\}$ bzw. $\gamma_0, \{\varphi_\lambda(x)\}$, erhält man genau die Behauptungen 1)*, 2)* (mit $c_\lambda=\frac{1}{\varrho_\lambda}\lim\limits_{x\to a+0}[g(x)\varphi'_\lambda(x)]$) oder 1**, 2)**, je nach dem Falle (I) bzw. (II).

2° Es sei $\{\bar{\varphi}_\lambda(x)\}$ ein Funktionensystem mit den Eigenschaften 1)* und 2)*.

Nun wendet man Satz 2, Teil I) mit $Q(x)=g(x)^{-1}$ an und erhält, daß 1. $\bar{\varphi}_\lambda(x)$ und zugleich $g(x)\bar{\varphi}'_\lambda(x)$ eine stetige Ableitung für $a<x<b$, $\lambda=1, 2, \ldots$ hat; 2. die Relationen

$$[g(x)\bar{\varphi}'_\lambda(x)]'=\bar{\varrho}_\lambda w(x)\bar{\varphi}_\lambda(x) \qquad (a<x<b;\lambda=1, 2, \ldots) \tag{4. 2}$$

und 3. die Randbedingungen

$$\bar{\varphi}_\lambda(a+0)=\bar{\varphi}_\lambda(b-0)=0, \quad \lim_{x\to a+0}[g(x)\bar{\varphi}'_\lambda(x)]=\bar{\varrho}_\lambda\bar{c}_\lambda \tag{4. 3}$$

mit

$$\bar{\varrho}_\lambda=-\left\{\int_a^b\frac{1}{g(x)}\left[\bar{\Phi}_\lambda(x)+\bar{c}_\lambda\right]^2dx\right\}^{-1}, \quad \bar{c}_\lambda=-\left[\int_a^b\frac{dx}{g(x)}\right]^{-1}\int_a^b\frac{1}{g(x)}\bar{\Phi}_\lambda(x)dx \tag{4. 4}$$
$$(\lambda=1, 2, \ldots)$$

befriedigt sind.

Daher schließt man aber wegen der $w$-Vollständigkeit von $\{\bar{\varphi}_\lambda(x)\}$ unmittelbar, daß $\{\bar{\varrho}_\lambda\}$ mit dem Spektrum $\{\varrho_\lambda\}$ zusammenfällt, und weiter, daß $\{\bar{\varphi}_\lambda(x)\}$, abgesehen von Vorzeichen und Reihenfolge, mit dem Systeme $\{\varphi_\lambda(x)\}$ identisch sein muß.

3° Falls $\{\bar{\varphi}_\lambda(x)\}$ die Eigenschaften 1)**, 2)** besitzt, so folgt aus Satz 2, Teil II) ähnlicherweise, daß $\{|\bar{\varphi}_\lambda(x)|\}$ nach einer geeigneten Umordnung in $\{|\varphi_\lambda(x)|\}$ übergeht, w. z. b. w.

**4. 2.** Satz 6 liefert z. B. eine Charakterisierung der trigonometrischen Systeme $\left\{\sqrt{\frac{2}{\pi}}\sin\lambda x\right\}$ und $\frac{1}{\sqrt{\pi}}, \left\{\sqrt{\frac{2}{\pi}}\cos\lambda x\right\}$ $(a=0, b=\pi, g(x)=w(x)\equiv 1)$, ferner des normierten Systems der Jacobischen Polynome $(a=-1, b=1, g(x)=$

$=(1-x)^{\alpha+1}(1+x)^{\beta+1}$, $w(x)=(1-x)^{\alpha}(1+x)^{\beta}$, $-1<\alpha<0$, $-1<\beta<0$). Das erste ist ein Beispiel für den Fall (4. 1) (I), das zweite und dritte entsprechen aber dem Falle (4. 1) (II).

Insbesondere ergibt sich: *für die Tschebyscheffschen Polynome erster Art* $T_\lambda(x)=\cos(\lambda \arccos x)$ $(\lambda=0,1,2,\ldots)$ *sind, abgesehen von konstanten Faktoren, folgende Eigenschaften charakteristisch:* 1) $T_\lambda(x)$ $(\lambda=0,1,2,\ldots)$ *sind in* $(-1,1)$ *stetig und bilden hier ein* $(1-x^2)^{-\frac{1}{2}}$*-vollständiges,* $(1-x^2)^{-\frac{1}{2}}$*-orthogonales System;* 2) $T_\lambda^*(x)=\int\limits_{-1}^{x}(1-u^2)^{-\frac{1}{2}}T_\lambda(u)\,du$ $(\lambda=1,2,\ldots)$ *ist* $(1-x^2)^{\frac{1}{2}}$*-orthogonal inbezug auf* $(-1,1)$.

Es sei noch, allgemeiner, *der Fall in* $[a,b]$ *stetiger und positiver* $g(x)$ *und* $w(x)$ hervorgehoben; dann bietet Satz 6, wie man sofort einsieht, *eine Charakterisierung jedes Systems von Sturm—Liouvilleschen Eigenfunktionen verknüpft mit*

$$[g(x)y']'=\varrho w(x)y \qquad (y=y(x);\ a\leqq x\leqq b)$$

*und den Randbedingungen* $y(a)=y(b)=0$ *oder* $y'(a)=y'(b)=0$.

**4. 3.** Die Funktionenfolge

$$1,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x,\ \ldots \tag{4.5}$$

hat die Eigenschaft, daß nicht nur sie selbst, sondern auch ihre erste und höhere Deriviertensysteme inbezug auf ein beliebiges Intervall von Länge $2\pi$ (zur Belegung 1) *orthogonal* sind; ein jedes dieser Systeme wird *vollständig*, wenn man 1 als Element hinzufügt. Wir wollen zeigen, daß diese Festsetzungen schon für (4.5) charakteristisch sind, ja sogar.

Satz 7. *Es seien* $\varphi_\lambda(x)$ $(\lambda=1,2,\ldots)$ *in einem endlichen Intervalle* $[a,b]$ *zweimal stetig differenzierbare Funktionen* $\neq 0$, *welche hier ein vollständiges Orthogonalsystem bilden und den Anfangsbedingungen* $\varphi_\lambda(a)\varphi_\lambda'(a)=0$ $(\lambda=1,2,\ldots)$ *genügen; es sei auch* $1,\{\varphi_\lambda'(x)\}$, *gleichwie* $1,\{\varphi_\lambda''(x)\}$ *orthogonal inbezug auf* $[a,b]$.[33]

*Dann stimmt* $\{\varphi_\lambda(x)\}$ *bis auf Reihenfolge und triviale lineare Transformationen mit* (4.5) *überein.*

Beweis. Unsere Annahmen implizieren

$$\left\{\begin{aligned} &\int_a^b \varphi_\lambda'(x)\,dx=0, \qquad \varphi_\lambda(a)=\varphi_\lambda(b)\\ &\int_a^b \varphi_\lambda''(x)\,dx=0, \qquad \varphi_\lambda'(a)=\varphi_\lambda'(b)\end{aligned}\right. \qquad (\lambda=1,2,\ldots) \tag{4.6}$$

[33] Statt der Orthogonalität von 1, $\{\varphi_\lambda'(x)\}$ und 1, $\{\varphi_\lambda''(x)\}$ mag nur diejenige von $\{\varphi_\lambda'(x)\}$ nebst den Randbedingungen $\varphi_\lambda(a)=\varphi_\lambda(b)$, $\varphi_\lambda'(a)=\varphi_\lambda'(b)$ $(\lambda=1,2,\ldots)$ verlangt werden. — In diesem Satze ist die Orthogonalität offenbar in engerem Sinne, d. h. mit der Belegung 1, zu verstehen.

und erlauben die Anwendung des Satzes 3 mit $g(x) = w(x) = 1$. Man sieht, daß die Bedingungen (2. 26) wegen (4. 6) erfüllt sind; somit ergibt die Orthogonalität von $\{\varphi_\lambda(x)\}$, $\{\varphi'_\lambda(x)\}$ inbezug auf $[a, b]$ und die Vollständigkeit des ersten Systems

(4. 7) $$\varphi''_\lambda(x) = \varrho_\lambda \varphi_\lambda(x) \qquad (a < x < b;\ \lambda = 1, 2, \ldots)$$

mit

(4. 8) $$\varrho_\lambda = -\left[\int_a^b \varphi'_\lambda(x)^2\, dx\right]\left[\int_a^b \varphi_\lambda(x)^2\, dx\right]^{-1} \qquad (\lambda = 1, 2, \ldots).$$

Durch die Transformation

(4. 9) $$t = \frac{2\pi}{b-a}(x-a), \quad x = a + \frac{b-a}{2\pi} t$$

ergibt es sich also: $y = \varphi_\lambda\left(a + \frac{b-a}{2\pi} t\right)$ $(\lambda = 1, 2, \ldots)$ genügen der Differentialgleichung

(4. 10) $$\frac{d^2y}{dt^2} = \varrho_\lambda \left(\frac{b-a}{2\pi}\right)^2 y \qquad (0 < t < 2\pi)$$

nebst den Bedingungen

(4. 11) $$y(0) = y(2\pi), \quad y'(0) = y'(2\pi).$$

Das Randwertproblem $\ddot{y} = \varrho y$ mit (4. 11) hat aber bekanntlich die Eigenwerte $\varrho = -n^2$ $(n = 0, 1, 2, \ldots)$, so daß

(4. 12) $$\varrho_\lambda \cdot \left(\frac{b-a}{2\pi}\right)^2 = -n^2, \quad \varrho_\lambda = -\left(\frac{2\pi}{b-a}\right)^2 n^2$$

für jede $\lambda = 1, 2, \ldots$ mit einem geeigneten ganzen $n$ gelten muß; jedem Eigenwerte gehören zwei linear unabhängige, bis auf konstante Faktoren bestimmte Eigenfunktionen: $\cos nt$, $\sin nt$ an.[34]

Wegen der Vollständigkeit von $\{\varphi_\lambda(x)\}$ inbezug auf $[a, b]$ und auf Grund der Bedingungen $\varphi_\lambda(a)\varphi'_\lambda(a) = 0$ $(\lambda = 1, 2, \ldots)$ muß die Folge $\left\{\varphi_\lambda\left(a + \frac{b-a}{2\pi} t\right)\right\}$, von konstanten Faktoren abgesehen, mitdem Systeme 1, $\cos t$, $\sin t$, $\cos 2t$, $\sin 2t, \ldots$ in $0 \leqq t \leqq 2\pi$ übereinstimmen; d. h. ist

$$\varphi_1(x) = C_1,\ \varphi_2(x) = C_2 \cos\left(\frac{2\pi}{b-a}(x-a)\right),\ \varphi_3(x) = C_3 \sin\left(\frac{2\pi}{b-a}(x-a)\right),$$

$$\varphi_4(x) = C_4 \cos\left(\frac{4\pi}{b-a}(x-a)\right),\ \varphi_5(x) = C_5 \sin\left(\frac{4\pi}{b-a}(x-a)\right), \ldots$$

für $a \leqq x \leqq b$, falls die Indices und die Konstanten $C$ passend gewählt sind.

Hiermit ist Satz 7 bewiesen.

[34] Vgl. z. B. [9].

**4. 4.** Betreffs *der klassischen orthogonalen Polynomsysteme* seien die folgenden Tatsachen hervorgehoben:[35] 1) Die Jacobischen Polynome $J_\lambda^{(\alpha,\beta)}(x) = \frac{(-1^\lambda)}{(2\lambda)!!}(1-x)^{-\alpha}(1+x)^{-\beta}\frac{d^\lambda}{dx^\lambda}(1-x)^{\alpha+\lambda}(1+x)^{\beta+\lambda}$ $(\alpha > -1, \beta > -1; \lambda = 0, 1, 2, \ldots)$, die (verallgemeinerten) Laguerreschen Polynome $\mathfrak{L}_\lambda^{(\alpha)}(x) = \frac{1}{\lambda!}x^{-\alpha}e^x\frac{d^\lambda}{dx^\lambda}(x^{\alpha+\lambda}e^{-x})$ $(\alpha > -1; \lambda = 0, 1, 2, \ldots)$ und die Hermiteschen Polynome $H_\lambda(x) = (-1)^\lambda e^{x^2}\frac{d^\lambda}{dx^\lambda}e^{-x^2}$ $(\lambda = 0, 1, 2, \ldots)$ sind, der Reihe nach, mit den folgenden Grundintervallen und Belegungsfunktionen verknüpft: $(-1, 1)$, $w(x) = (1-x)^\alpha(1+x)^\beta$; $(0, \infty)$, $w(x) = x^\alpha e^{-x}$; bzw. $(-\infty, \infty)$, $w(x) = e^{-x^2}$. Alle drei Systeme sind inbezug auf das zugehörige Intervall und $w(x)$ vollständig. 2) Es gelten die Relationen:

$$\frac{d}{dx}J_\lambda^{(\alpha,\beta)}(x) = \frac{1}{2}(\lambda+\alpha+\beta+1)J_{\lambda-1}^{(\alpha+1,\beta+1)}(x), \quad \frac{d}{dx}\mathfrak{L}_\lambda^{(\alpha)}(x) = -\mathfrak{L}_{\lambda-1}^{(\alpha+1)}(x),$$
$$\frac{d}{dx}H_\lambda(x) = 2\lambda H_{\lambda-1}(x) \qquad (\lambda = 1, 2, \ldots).$$

Wie daraus erhellt, sind auch die ersten, zweiten usw. Ableitungen eines genannten Systems *vollständig* und *orthogonal* inbezug auf dasselbe Intervall, und zwar, der Reihe nach, zu solchen Belegungen, die (samt $w(x)$) eine geometrische Progression bilden und in den Endpunkten des Grundgebietes verschwinden, sonst aber positiv bleiben. Unser nächstes Ziel ist zu zeigen, daß diese Eigenschaften, erweitert mit gewissen natürlichen Bedingungen über Beschränktheit, schon *charakteristisch* sind; ja sogar, es genügt im wesentlichen *die ersten drei (und nicht wenigere) Ableitungssysteme* zu betrachten. Schärfer und genauer:

Satz 8. *Es seien $\varphi_\lambda(x)$ $(\lambda = 1, 2, \ldots)$ in $(a, b)$ viermal stetig differenzierbare Funktionen, welche darauf ein w-vollständiges, w-orthogonales System bilden. Wir nehmen an, 1) $\{\varphi_\lambda'(x)\}$, $\{\varphi_\lambda''(x)\}$, $\{\varphi_\lambda'''(x)\}$ seien auch vollständig und orthogonal inbezug auf $(a, b)$ und gewisse Belegungsfunktionen $g_1(x)$, $g_2(x)$ bzw. $g_3(x)$; 2) $w(x)$, $g_1(x)$, $g_2(x)$, $g_3(x)$ seien in $(a, b)$ positiv, zweimal stetig differenzierbar und bilden eine geometrische Progression; 3) Sei $w(x) \in L(a, b)$, $g_1(a+0) = g_1(b-0) = 0$, während $[g_1(x)w(x)^{-1}]'$, $g_1'(x)w(x)^{-1}$, gleichwie jedes $\varphi_\lambda(x)$ und ihre genannten Ableitungen in einem beliebigen endlichen Intervalle $[x_1, x_2] \in [a, b]$ beschränkt bleiben sollen;*[36] *schließlich, wenn $[a, b]$ ins Unendliche reicht, 4) sei $w(x) = o(|x|^{-M})$ für jedes $M > 0$, während $[g_1(x)w(x)^{-1}]'$, $g_1'(x)w(x)^{-1}$, ferner alle $\varphi_\lambda(x)$ und ihre genannten Ableitungen für $|x| \to \infty$ höchstens wie eine Potenz von $|x|$ unendlich werden sollen.*

[35] Vgl. z. B. [29].

[36] Ist $[a, b]$ endlich, so mag der Ausdruck: „in einem beliebigen endlichen Intervalle $[x_1, x_2] \in [a, b]$" einfach mit „in $[a, b]$" ersetzt werden.

*Dann stimmt $\{\varphi_\lambda(x)\}$ bis auf Reihenfolge und triviale lineare Transformationen mit einem der klassischen Polynomsysteme überein, und zwar mit demjenigen von Jacobi, Laguerre oder Hermite, je nachdem $[a, b]$ endlich ist, bzw. auf einer Seite oder beiden Seiten ins Unendliche reicht. — Zugleich müssen $w(x)$ und $g_1(x)$, abgesehen von geeigneten linearen Transformationen, mit den entsprechenden Gewichtsfunktionen identisch sein.*

BEWEIS. Es sei angenommen, daß alle genannten Bedingungen befriedigt sind.

1° In einem beliebigen endlichen und abgeschlossenen Intervalle, welches $[a, b]$ angehört, bleibt der Quotient $g_1(x)w(x)^{-1}$ wegen

$$(4.13)\quad \frac{g_1(x)}{w(x)}=\frac{g_1(x_0)}{w(x_0)}+\int_{x_0}^{x}\left[\frac{g_1(u)}{w(u)}\right]' du \leqq \frac{g_1(x_0)}{w(x_0)}+(B-A)\max_{A\leqq u\leqq B}\left|\left[\frac{g_1(u)}{w(u)}\right]'\right|$$
$$(a<A\leqq x_0<x\leqq B<b)$$

beschränkt (vgl. 3)).

Im Falle eines unendlichen $(a, b)$ hat man (vgl. 4))

$$(4.14)\quad \left[\frac{g_1(x)}{w(x)}\right]'=O(|x|^\mu)$$

mit einem geeigneten $\mu\geqq 0$; dies ergibt die Abschätzung

$$\frac{g_1(x)}{w(x)}\leqq\frac{g_1(x_0)}{w(x_0)}+K\int_{x_0}^{x}|u|^\mu du+K_1(x-x_0)$$

mit $K>0$, $K_1>0$ und für jedes $x\in(a, b)$, $x_0\in(a, b)$, so daß man

$$(4.15)\quad \frac{g_1(x)}{w(x)}=O(|x|^{\mu+1})$$

erhält.

Die Bedingung 2) liefert

$$(4.16)\quad g_2(x)=\frac{g_1(x)^2}{w(x)},\quad g_3(x)=\frac{g_1(x)^3}{w(x)^2};$$

daher folgt wegen (4.13) und 3)

$$(4.17)\quad \lim g_2(x)=0,\quad \lim g_3(x)=0$$

in jedem Endpunkte von $(a, b)$, welcher im Endlichen liegt.

Ist $(a, b)$ unendlich, so implizieren die Bedingungen 3), 4) und (4.15), (4.16)

$$(4.18)\quad \left.\begin{matrix} g_1(x)\\ g_2(x)\\ g_3(x)\end{matrix}\right\}=o(|x|^{-N})\qquad (N>0 \text{ beliebig}).$$

Wir erwähnen ferner die aus (4. 16) fließenden Formeln

$$\left.\begin{aligned}&(4.\,19)\qquad \frac{g_2(x)}{g_1(x)}=\frac{g_1(x)}{w(x)},\quad \frac{g_2'(x)}{g_1(x)}=\frac{g_1'(x)}{w(x)}+\left[\frac{g_1(x)}{w(x)}\right]'\\ &(4.\,20)\qquad \frac{g_3(x)}{g_2(x)}=\frac{g_1(x)}{w(x)},\quad \frac{g_3'(x)}{g_2(x)}=\frac{g_1'(x)}{w(x)}+2\left[\frac{g_1(x)}{w(x)}\right]'\end{aligned}\right\}\quad (a<x<b),$$

welche samt den vorigen zeigen, daß die Funktionen $g_2(x)g_1(x)^{-1}$, $g_2'(x)g_1(x)^{-1}$ und $g_3(x)g_2(x)^{-1}$, $g_3'(x)g_2(x)^{-1}$ in jedem endlichen abgeschlossenen Intervalle in $[a, b]$ beschränkt sind, während im Falle eines unendlichen $(a, b)$ für $|x|\to\infty$ höchstens wie eine Potenz von $|x|$ unendlich werden können.

2° Wegen unserer Voraussetzungen ist Satz 3 anwendbar. Z. B. gilt $w(x)^{-1}[g_1(x)\varphi_\lambda'(x)]'\in L^2_w(a, b)$, da

$$\frac{1}{\sqrt{w(x)}}[g_1(x)\varphi_\lambda'(x)]'=\sqrt{w(x)}\left[\frac{g_1'(x)}{w(x)}\varphi_\lambda'(x)+\frac{g_1(x)}{w(x)}\varphi_\lambda''(x)\right]\quad (a<x<b),$$

und der Ausdruck rechter Hand, laut der Bedingungen 1)—4) und 1°, zur Klasse $L^2(a, b)$ gehört; es folgt ähnlicherweise aus 3), 4) und (4. 18)

$$\lim_{x\to a+0}[g_1(x)\varphi_m(x)\varphi_n'(x)]=\lim_{x\to b-0}[g_1(x)\varphi_m(x)\varphi_n'(x)]=0\qquad (m, n=1, 2, \ldots).$$

Also ergeben die $w$-Orthogonalität, $w$-Vollständigkeit von $\{\varphi_\lambda(x)\}$ und die $g_1$-Orthogonalität von $\{\varphi_\lambda'(x)\}$ die Differentialgleichungen

$$(4.\,21)\qquad [g_1(x)\varphi_\lambda'(x)]'=\varrho_\lambda w(x)\varphi_\lambda(x)\qquad (a<x<b;\ \lambda=1, 2, \ldots)$$

mit

$$(4.\,22)\qquad \varrho_\lambda=-\left[\int_a^b g_1(x)\varphi_\lambda'(x)^2\,dx\right]\left[\int_a^b w(x)\varphi_\lambda(x)^2\,dx\right]^{-1}\qquad (\lambda=1, 2, \ldots).$$

Nach einer geeigneten Abänderung der Bezeichnungen und durch Benützung von (4. 19), (4. 20) schließt man aus der $g_1$-Orthogonalität, $g_1$-Vollständigkeit von $\{\varphi_\lambda'(x)\}$ und der $g_2$-Orthogonalität von $\{\varphi_\lambda''(x)\}$

$$(4.\,23)\qquad [g_2(x)\varphi_\lambda''(x)]'=\varrho_\lambda^* g_1(x)\varphi_\lambda'(x)\qquad (a<x<b;\ \lambda=1, 2, \ldots)$$

mit

$$(4.\,24)\qquad \varrho_\lambda^*=-\left[\int_a^b g_2(x)\varphi_\lambda''(x)^2\,dx\right]\left[\int_a^b g_1(x)\varphi_\lambda'(x)^2\,dx\right]^{-1}\quad (\lambda=1, 2, \ldots),$$

während die $g_2$-Orthogonalität, $g_2$-Vollständigkeit von $\{\varphi_\lambda''(x)\}$ und die $g_3$-Orthogonalität von $\{\varphi_\lambda'''(x)\}$ zu den Relationen

$$(4.\,25)\qquad [g_3(x)\varphi_\lambda'''(x)]'=\varrho_\lambda^{**} g_2(x)\varphi_\lambda''(x)\qquad (a<x<b;\ \lambda=1, 2, \ldots)$$

mit

$$(4.\,26)\qquad \varrho_\lambda^{**}=-\left[\int_a^b g_3(x)\varphi_\lambda'''(x)^2\,dx\right]\left[\int_a^b g_2(x)\varphi_\lambda''(x)^2\,dx\right]^{-1}\quad (\lambda=1, 2, \ldots)$$

führen.

3° Auf Grund von (4. 21), (4. 23), (4. 25) und der oben benutzten Randbedingungen ist leicht zu erkennen, daß $\{\varphi_\lambda(x)\}$, $\{\varphi_\lambda'(x)\}$, $\{\varphi_\lambda''(x)\}$, nach

zufälliger Weglassung gewisser Elemente, *gewöhnliche SL-Systeme bilden* welche der Reihe nach mit den Eigenwerten $\{\varrho_\lambda\}$, $\{\varrho_\lambda^*\}$, $\{\varrho_\lambda^{**}\}$ und den Belegungsfunktionen $(g_1(x), w(x))$, $(g_2(x), g_1(x))$ bzw. $(g_3(x), g_2(x))$ verknüpft sind; dann hat man aber, laut Satz 5, Teil II), u. a. die Relationen

$$\left.\begin{aligned} &\frac{g_1(x)}{w(x)}=px^2+qx+r && (4.27)\\ &\frac{g_1'(x)}{w(x)}=Px+Q && (4.28)\end{aligned}\right\}\qquad (a<x<b)$$

mit konstanten $p, q, r, P$ und $Q$.

Die Koeffizienten $p, q, r$ können, wegen der Positivität von $g_1(x)$ in $(a, b)$, offenbar nicht gleichzeitig verschwinden, d. h.

$$p^2+q^2+r^2>0; \tag{4.29}$$

ferner muß

$$P\neq 0 \tag{4.30}$$

sein, weil man sonst durch die Randbedingungen 3)

$$\lim_{x\to b-0} g_1(x)=Q\int_a^b w(t)\,dt=0,$$

also (vgl. (4.28)) $g_1(x)\equiv 0$ folgern könnte, in Gegensatz zu 2).

Aus (4.27) und (4.28) resultiert

$$\frac{g_1'(x)}{g_1(x)}=\frac{Px+Q}{px^2+qx+r}, \tag{4.31}$$

d. h. $y=g_1(x)$ genügt notwendigerweise der sogenannten *Pearsonschen Differentialgleichung*[37]

$$y'=\frac{Px+Q}{px^2+qx+r}\,y. \tag{4.32}$$

Setzen wir die Ausdrücke von $g_1(x)$ und $g_1'(x)$ aus (4.27) und (4.28) in (4.21), so folgt

$$(px^2+qx+r)\varphi_\lambda''(x)+(Px+Q)\varphi_\lambda'(x)-\varrho_\lambda\varphi_\lambda(x)=0 \tag{4.33}$$
$$(a<x<b;\ \lambda=1,2,\ldots),$$

d. h. $y=\varphi_\lambda(x)$ soll eine Lösung *der Differentialgleichung vom Fuchsschen Typus*[38]

$$(px^2+qx+r)y''+(Px+Q)y-\varrho_\lambda y=0 \qquad (a<x<b) \tag{4.34}$$

sein.

4° Nun gebrauchen wir folgendes

Lemma 2.[39] *Es sei $y=y(x)$ eine für $a<x<b$ stetig differenzierbare und nicht verschwindende Funktion, welche hier* (4.32) *nebst den Randbedin-*

[37] Im Falle der sogenannten *Pearsonschen Wahrscheinlichkeitsverteilungen* genügt die Verteilungsfunktion bekanntlich einer Differentialgleichung von der Form (4.32). Vgl. [3].

[38] Vgl. z. B. [19], S. 131.

[39] Vgl. z. B. [18], Ch. VI, VII.

*gungen* $y(a+0)=y(b-0)=0$ *genügt; im Falle eines unendlichen* $(a,b)$ *bestehe überdies* $y(x)=o(|x|^{-N})$ *für jedes* $N>0$. *Dann hat man (mit geeignetem konstantem* $C\neq 0$)

1) $p\neq 0$, $px^2+qx+r\equiv p(x-a)(x-b)$ *und für* $a<x<b$

$$y(x)=C(b-x)^{\varkappa_1}(x-a)^{\varkappa_2} \tag{4.35}$$

*mit*

$$\varkappa_{1,2}=\frac{P}{2p}\pm\frac{1}{p(b-a)}\left(Q-\frac{Pq}{2p}\right)>0, \tag{4.36}$$

*falls* $[a,b]$ *endlich ist;*

2) $p=0$, $Pq<0$, $a=-\frac{r}{q}$ *und für* $a<x<\infty$

$$y(x)=C(x-a)^{\tau}e^{\frac{P}{q}x} \tag{4.37}$$

*mit*

$$\tau=\frac{1}{q}\left(Q-\frac{Pr}{q}\right)>0, \tag{4.38}$$

*wenn a endlich ist und* $b=\infty$;

3) $p=0$, $Pq>0$, $b=-\frac{r}{q}$ *und für* $-\infty<x<b$

$$y(x)=C(b-x)^{\tau}e^{\frac{P}{q}x} \tag{4.39}$$

*mit* (4.38), *wenn* $a=-\infty$ *und b endlich ist;*

4) $p=q=0$, $Pr<0$ *und für* $-\infty<x<\infty$

$$y(x)=C\exp\left(\frac{P}{2r}x^2+\frac{Q}{r}x\right), \tag{4.40}$$

*falls* $a=-\infty$, $b=\infty$.

Denn (4.32) läßt sich sofort durch logarithmische Integration lösen; mit Rücksicht auf (4.29) und (4.30) hat man die folgenden Fälle zu unterscheiden (wobei $x$ ein Stetigkeitsintervall des rationalen Koeffizienten rechter Hand in (4.32) durchläuft, die Konstante $C$ aber beliebig gewählt werden mag):

1) ist $p\neq 0$, $q^2>4pr$, so erhält man

$$y=C|x-\xi_1|^{\varkappa_1}|x-\xi_2|^{\varkappa_2} \tag{4.41}$$

mit

$$\xi_{1,2}=\frac{1}{2p}\left(-q\pm\sqrt{q^2-4pr}\right),\quad \varkappa_{1,2}=\frac{P}{2p}\pm\left(Q-\frac{Pq}{2p}\right)\frac{1}{p(\xi_1-\xi_2)};$$

2) ist $p \neq 0$, $q^2 = 4pr$, so ergibt sich

$$y = C\left|x + \frac{q}{2p}\right|^{\frac{P}{q}} \exp\left[\left(\frac{Pq}{p} - 2Q\right)\frac{1}{2px+q}\right]; \tag{4.42}$$

3) ist $p \neq 0$, $q^2 < 4pr$, so resultiert

$$y = C|px^2 + qx + r|^{\frac{P}{2p}} \exp\left[\left(2Q - \frac{Pq}{p}\right)\frac{1}{\sqrt{4pr - q^2}} \operatorname{arc\,tg} \frac{2px+q}{\sqrt{4pr-q^2}}\right]; \tag{4.43}$$

4) im Falle $p = 0$, $q \neq 0$ folgt

$$y = Ce^{\frac{P}{q}x}\left|x + \frac{r}{q}\right|^{\tau} \tag{4.44}$$

mit $\tau = \frac{1}{q}\left(Q - \frac{Pr}{q}\right)$;

5) im Falle $p = q = 0$ ist

$$y = C\exp\left(\frac{P}{2r}x^2 + \frac{Q}{r}x\right). \tag{4.45}$$

Von (4.42)—(4.45) liest man leicht ab, daß die Annahmen unseres Lemmas und die Fälle 2), 3) unverträglich sind, und weiter, daß diese die Gültigkeit einer von den fraglichen Behauptungen 1)—4) nach sich ziehen, je nachdem $[a, b]$ endlich, einerseits bzw. beiderseits unendlich ist.[40]

5° Nun kehren wir auf (4.31) und (4.33) zurück.

Ist $[a, b]$ endlich, so hat man nach (4.31) und dem eben bewiesenen Lemma die Behauptungen 1) mit $y(x) = g_1(x)$ und $C > 0$; wir setzen

$$\varkappa_1 = \alpha + 1, \quad \varkappa_2 = \beta + 1, \tag{4.46}$$

so daß (vgl. (4.36)) $\alpha > -1$, $\beta > -1$. Also ist

$$g_1(x) = C(b-x)^{\alpha+1}(x-a)^{\beta+1} \qquad (a < x < b;\ \alpha > -1,\ \beta > -1,\ C > 0) \tag{4.47}$$

und wegen (4.27)

$$w(x) = -\frac{C}{p}(b-x)^{\alpha}(x-a)^{\beta} \tag{4.48}$$
$$(a < x < b;\ \alpha > -1,\ \beta > -1,\ C > 0,\ p < 0).$$

Andererseits, falls man den Teil 1) des Lemmas und die lineare Transformation

$$x = a + \frac{b-a}{2}(t+1), \quad t = -1 + \frac{2}{b-a}(x-a)$$

[40] Die allgemeinen Lösungen (4.41)—(4.45) liefern unmittelbar eine partikuläre Integralkurve von (4.32), welche durch einen gegebenen nicht-singulären Punkt der $xy$-Ebene geht.

auf (4.33) anwendet, so folgt (vgl. (4.46))

$$
\begin{aligned}
&(1-t^2)\frac{d^2}{dt^2}\varphi_\lambda\left(a+\frac{b-a}{2}(t+1)\right)-\\
(4.49)\qquad &-[(\alpha+\beta+2)t+(\alpha-\beta)]\frac{d}{dt}\varphi_\lambda\left(a+\frac{b-a}{2}(t+1)\right)+\\
&+\frac{\varrho_\lambda}{p}\varphi_\lambda\left(a+\frac{b-a}{2}(t+1)\right)=0 \quad (-1<t<1;\ \alpha>-1,\ \beta>-1);
\end{aligned}
$$

d. h. $y=\varphi_\lambda\left(a+\frac{b-a}{2}(t+1)\right)$ genügt der Differentialgleichung

$$(4.50)\qquad (1-t^2)\frac{d^2y}{dt^2}-[(\alpha+\beta+2)t+(\alpha-\beta)]\frac{dy}{dt}+Ky=0$$

oder (in Sturm—Liouvillescher Form geschrieben)

$$(4.51)\qquad \frac{d}{dt}\left[(1-t)^{\alpha+1}(1+t)^{\beta+1}\frac{dy}{dt}\right]=-K(1-t)^\alpha(1+t)^\beta$$

mit $K=\frac{\varrho_\lambda}{p}$ für $-1<t<1$ und $\lambda=1,2,\ldots$ .

Aber das Randwertproblem: „gesucht die für $-1<t<1$ beschränkten Lösungen von (4.50) oder (4.51)" hat bekanntlich die (einfachen) Eigenwerte $K=-(\lambda-1)(\lambda+\alpha+\beta)$ $(\lambda=1,2,\ldots)$ und die Eigenfunktionen const$\cdot$ $\cdot J_{\lambda-1}^{(\alpha,\beta)}(t)$ $(\lambda=1,2,\ldots;$[41] da unsere Voraussetzungen 3) die Beschränktheit alle $\varphi_\lambda\left(a+\frac{b-a}{2}(t+1)\right)$ in $[-1,1]$ implizieren, ergibt sich bei einer geeigneten Wahl der Indices

$$(4.52)\qquad \varrho_\lambda=pK=p(\lambda-1)(\lambda+\alpha+\beta) \qquad (\lambda=1,2,\ldots)$$

und (wegen der angenommenen Vollständigkeit)

$$(4.53)\qquad \varphi_\lambda\left(a+\frac{b-a}{2}(t+1)\right)=\text{const}\cdot J_{\lambda-1}^{(\alpha,\beta)}(t) \qquad (-1<t<1;\ \lambda=1,2,\ldots),$$

d. h.

$$(4.54)\qquad \varphi_\lambda(x)=\text{const}\cdot J_{\lambda-1}^{(\alpha,\beta)}\left(\frac{2}{p-a}(x-a)-1\right) \qquad (a<x<b;\ \lambda=1,2,\ldots)$$

mit $\alpha>-1$, $\beta>-1$.

6° Ist $a$ endlich, $b=\infty$, so ergibt die Anwendung des Teils 2) unseres Lemmas mit $y(x)=g_1(x)$, daß $p=0$, $Pq<0$, $a=-\frac{r}{q}$ und für $a<x<\infty$

$$(4.55)\qquad g_1(x)=C(x-a)^{\alpha+1}e^{\frac{P}{q}x}$$

[41] Vgl. z. B. [19], S. 465 oder [4], S. 258—260; unsere Bezeichnungen sind diejenigen von [29]. — Die betreffende Behauptung kann am kürzesten aus der Vollständigkeit des Jacobischen Polynomsystems inbezug auf $[-1,1]$ und die entsprechenden Gewichtsfunktionen gefolgert werden. ([29], S. 39.)

mit $C>0$, $\alpha+1=\tau=\frac{1}{q}\left(Q-\frac{Pr}{q}\right)$ und (vgl. (4.38))

(4.56) $$\alpha>-1.$$

Kombiniert man (4.55) mit (4.27), so folgt noch

(4.57) $$w(x)=\frac{C}{q}(x-a)^{\alpha}e^{\frac{P}{q}x}$$

mit $q>0$. (4.33) kann nun in der Form

(4.58) $$q(x-a)\varphi_\lambda''(x)+(Px+Q)\varphi_\lambda'(x)-\varrho_\lambda\varphi_\lambda(x)=0 \qquad (a<x<\infty)$$

geschrieben werden, welche sich durch

(4.59) $$x=a-\frac{q}{P}t, \qquad t=\frac{P}{q}(a-x)$$

in die Relation

(4.60) $$t\frac{d^2}{dt^2}\varphi_\lambda\left(a-\frac{q}{P}t\right)+[(\alpha+1)-t]\frac{d}{dt}\varphi_\lambda\left(a-\frac{q}{P}t\right)+ \\ +\frac{\varrho_\lambda}{P}\varphi_\lambda\left(a-\frac{q}{P}t\right)=0 \qquad (0<t<\infty;\ \alpha>-1)$$

transformiert; d. h. $y=\varphi_\lambda\left(a-\frac{q}{P}t\right)$ genügt der Differentialgleichung

(4.61) $$t\frac{d^2y}{dt^2}+[(\alpha+1)-t]\frac{dy}{dt}+Ky=0$$

oder

(4.62) $$\frac{d}{dt}\left(t^{\alpha+1}e^{-t}\frac{dy}{dt}\right)=-Kt^{\alpha}e^{-t}y$$

mit $K=\frac{\varrho_\lambda}{P}$ für $0<t<\infty$ und $\lambda=1,2,\ldots$ .

Das durch (4.61) bzw. (4.62) und die Bedingungen: „$y=y(t)$ sei beschränkt für $t\to+0$ und $O(t^{\nu})$ mit geeignetem $\nu\geqq 0$ für $t\to+\infty$“ gegebene Randwertproblem hat aber die (einfachen) Eigenwerte $K=\lambda-1$ $(\lambda=1,2,\ldots)$ und die Eigenfunktionen const$\cdot\mathfrak{L}_{\lambda-1}^{(\alpha)}(t)$ $(\lambda=1,2,\ldots)$;[42] da die Voraussetzungen 4) $\varphi_\lambda\left(a-\frac{q}{P}t\right)=O(t^{\nu})$ $(0\leqq t<\infty)$ implizieren, gelangt man nach einer geeigneten Umordnung zu

(4.63) $$\varrho_\lambda=PK=P(\lambda-1) \qquad (\lambda=1,2,\ldots),$$

[42] Vgl. z. B. [19], S. 429 und [4], S. 261. — Das ist eine einfache Konsequenz der Vollständigkeit von $\{\mathfrak{L}_{\lambda}^{(\alpha)}(t)\}$ inbezug auf $(0,\infty)$ und $w(t)=t^{\alpha}e^{-t}$ $(\alpha>-1)$ ([29], S. 104).

und (wegen der Vollständigkeit)

$$\varphi_\lambda\left(a-\frac{q}{P}t\right)=\text{const}\cdot\mathfrak{L}_{\lambda-1}^{(\alpha)}(t) \qquad (0<t<\infty;\ \lambda=1,2,\ldots), \tag{4.64}$$

d. h.

$$\varphi_\lambda(x)=\text{const}\cdot\mathfrak{L}_{\lambda-1}^{(\alpha)}\left(\frac{P}{q}(a-x)\right) \qquad (a<x<\infty;\ \lambda=1,2,\ldots). \tag{4.65}$$

7° Falls $a=-\infty$ und $b$ endlich ist, bedarf man nur einer einfachen Veränderung in den Bezeichnungen von 6°. Wir erhalten

$$\left.\begin{aligned} &g_1(x)=C(b-x)^{\alpha+1}e^{\frac{P}{q}x} &(4.66)\\ &w(x)=-\frac{C}{q}(b-x)^{\alpha}e^{\frac{P}{q}x} &(4.67)\end{aligned}\right\} \quad (C>0,\ q<0,\ P<0,\ \alpha>-1),$$

$$\varphi_\lambda(x)=\text{const}\cdot\mathfrak{L}_{\lambda-1}^{(\alpha)}\left(\frac{P}{q}(b-x)\right) \qquad (-\infty<x<b;\ \alpha>-1,\ \lambda=1,2,\ldots). \tag{4.68}$$

8° Schließlich, im Falle $a=-\infty$, $b=\infty$ kann Teil 4) des Lemmas benutzt werden. Bezüglich (4.31) bekommt man $p=q=0$, $Pr<0$ und für $-\infty<x<\infty$

$$g_1(x)=C\exp\left(\frac{P}{2r}x^2+\frac{Q}{r}x\right); \tag{4.69}$$

daher und aus (4.27)

$$w(x)=\frac{C}{r}\exp\left(\frac{P}{2r}x^2+\frac{Q}{r}x\right) \tag{4.70}$$

mit $r>0$.

Andererseits reduziert (4.27) sich nun auf

$$r\varphi_\lambda''(x)+(Px+Q)\varphi_\lambda'(x)-\varrho_\lambda\varphi_\lambda(x)=0 \qquad (-\infty<x<\infty), \tag{4.71}$$

was durch

$$x=t\sqrt{-\frac{2r}{P}}-\frac{Q}{P},\quad t=x\sqrt{-\frac{P}{2r}}+\frac{Q}{\sqrt{-2Pr}} \tag{4.72}$$

in die Relation

$$\begin{aligned}&\frac{d^2}{dt^2}\varphi_\lambda\left(t\sqrt{-\frac{2r}{P}}-\frac{Q}{P}\right)-2t\frac{d}{dt}\varphi_\lambda\left(t\sqrt{-\frac{2r}{P}}-\frac{Q}{P}\right)+\\ &\qquad+\frac{2\varrho_\lambda}{P}\varphi_\lambda\left(t\sqrt{-\frac{2r}{P}}-\frac{Q}{P}\right)=0 \qquad (-\infty<t<\infty)\end{aligned} \tag{4.73}$$

übergeht.

Also sind $y=\varphi_\lambda\left(t\sqrt{-\frac{2r}{P}}-\frac{Q}{P}\right)$ $(\lambda=1,2,\ldots)$ Lösungen von

(4.74) $$\frac{d^2y}{dt^2}-2t\frac{dy}{dt}+Ky=0,$$

(4.75) $$\frac{d}{dt}\left(e^{-t^2}\frac{dy}{dt}\right)=-Ke^{-t^2}y$$

mit $K=\frac{2\varrho_\lambda}{P}$ für $-\infty<t<\infty$.

Das Problem: „gesucht sind Funktionen $y=y(t)$, welche (4.74) bzw. (4.75) und den Randbedingungen $y(t)=O(|t|^\mu)$ (mit einem geeigneten $\mu\geqq 0$) für $t\to\pm\infty$ genügen" — hat, wie bekannt, die (einfachen) Eigenwerte $K=2(\lambda-1)$ $(\lambda=1,2,\ldots)$ und die Eigenfunktionen const$\cdot H_{\lambda-1}(t)$ $(\lambda=1,2,\ldots)$;[43] wegen unserer Voraussetzungen 4) folgt also aus (4.75)

(4.76) $$\varrho_\lambda=\frac{P}{2}K=P(\lambda-1) \qquad (\lambda=1,2,\ldots),$$

ferner

(4.77) $$\varphi_\lambda\left(t\sqrt{-\frac{2r}{P}}-\frac{Q}{P}\right)=\text{const}\cdot H_\lambda(t) \qquad (-\infty<t<\infty;\ \lambda=1,2,\ldots),$$

d. h.

(4.78) $$\varphi_\lambda(x)=\text{const}\cdot H_\lambda\left(x\sqrt{-\frac{P}{2r}}+\frac{Q}{\sqrt{-2Pr}}\right) \quad (-\infty<x<\infty;\ \lambda=1,2,\ldots).$$

Somit ist der Beweis vollendet.

## § 5. Verallgemeinerung einiger Sätze über die Differentiation gewöhnlicher Fourierreihen

**5. 1.** Wie bekannt, darf man eine trigonometrische Fourierreihe stets gliedweise integrieren, dagegen nur ausnahmsweise formal differenzieren: die Größenordnung der Fourierkoeffizienten wird im ersten Falle verbessert, im zweiten aber verschlimmert. Im Anschluß einer Fragestellung von ABEL hat jedoch L. FEJÉR bewiesen, daß die gliedweise differenzierte Fourierreihe einer Funktion $f(x)$ $(0\leqq x\leqq 2\pi)$ im Inneren des Grundintervalles $C_1$-summierbar ist und $f'(x)$ zur $C_1$-Summe hat, falls diese Ableitung z. B. in $(0, 2\pi)$ existiert und stetig ist.[44]

[43] Vgl. z. B. [19], S. 415 und [4], S. 261; für die Vollständigkeit von $\{H_\lambda(t)\}$ s. auch [29], S. 104.

[44] [6], [7]. Die stetige Differenzierbarkeit von $f(x)$ impliziert natürlich die Konvergenz der Fourierreihe selbst. — Spätere Arbeiten von YOUNG, DE LA VALLÉE POUSSIN, GRONWALL, ZYGMUND ([32]) haben sich mit gewissen Verschärfungen und mit dem Falle höherer Ableitungen beschäftigt.

Die Ergebnisse der vorigen Abschnitte ergeben nun einen allgemeinen Satz über Konvergenz und Differentiation solcher Orthogonalreihen, welche eine orthogonale Ableitung haben.

SATZ 9. *Es seien* $w(x)$ *und* $g(x)$ *in einem endlichen* $[a, b]$ *positive, zweimal stetig differenzierbare Gewichtsfunktionen.* $\{\varphi_\lambda(x)\}$ *sei ein w-vollständiges, w-orthonormales System ebenda mit stetigen dritten Ableitungen versehener Funktionen mit* $\varphi_\lambda(a)=\varphi_\lambda(b)=0$ $(\lambda=1,2,\ldots)$ *und* $\{\varphi'_\lambda(x)\}$ *sei g-orthogonal in* $(a, b)$*; bezüglich der Reihenfolge nehmen wir an, daß die g-Norm von* $\varphi'_\lambda(x)$ $(\lambda=1,2,\ldots)$ *zunimmt. — Es bezeichnen* $\gamma_\lambda$ $(\lambda=1,2,\ldots)$ *die Fourierkoeffizienten einer Funktion* $f(x)\in \mathrm{Lip}\, 1$ $(a\leqq x\leqq b)$[45] *inbezug auf* $\varphi_\lambda(x)$ $(\lambda=1,2,\ldots)$.

*Dann ist die Fourierreihe* $\sum_{\lambda=1}^{\infty}\gamma_\lambda\varphi'_\lambda(x)$ *in* $(a, b)$ *konvergent und sie stellt dort in jedem inneren Teilintervalle gleichmäßig* $f(x)$ *dar, während* $\sum_{\lambda=1}^{\infty}\gamma_\lambda\varphi'_\lambda(x)$ *für fast jedes* $x\in(a,b)$ $C_1$*-summierbar ist und* $f'(x)$ *zur* $C_1$*-Summe hat. Letzteres gilt gewiß in solchen Punkten von* $(a, b)$, *wo* $f'(x)$ *existiert und stetig ist, und die Summierbarkeit ist in jedem abgeschlossenen inneren Stetigkeitsintervall der Ableitung* $f'(x)$ *eine gleichmäßige.*

Der BEWEIS stützt sich u. a. auf wohlbekannte Tatsachen über *Sturm—Liouvillesche Orthogonalsysteme*, insbesondere auf die asymptotische Formel von HOBSON und den *Haarschen Äquikonvergenzsatz;* wir brauchen noch eine Eigenschaft der auftretenden Kernfunktionen, nämlich die Tatsache, daß sie in den Endpunkten zum Werte Null $C_1$-limitierbar sind.

1° Die vorausgesetzten Orthogonalitätseigenschaften von $\{\varphi_\lambda(x)\}$, $\{\varphi'_\lambda(x)\}$ haben nach dem Satze 3 zur Folge, daß die Relationen

$$[g(x)\varphi'_\lambda(x)]'=\varrho_\lambda w(x)\varphi_\lambda(x)\qquad(a\leqq x\leqq b;\ \lambda=1,2,\ldots)\tag{5.1}$$

mit

$$\left.\begin{aligned}\varrho_\lambda&=-\int_a^b g(x)\varphi'_\lambda(x)^2dx &&(5.2)\\ \varphi_\lambda(a)&=\varphi_\lambda(b)=0 &&(5.3)\end{aligned}\right\}\qquad(\lambda=1,2,\ldots)$$

gelten; wegen der $w$-Vollständigkeit von $\{\varphi_\lambda(x)\}$ sind also $y=\mathrm{const}\cdot\varphi_\lambda(x)$ $(\lambda=1,2,\ldots)$ *alle* Eigenfunktionen des Sturm—Liouvilleschen Randwertproblems

$$[g(x)y']'=\varrho w(x)y\qquad(a\leqq x\leqq b);\quad y(a)=y(b)=0\tag{5.4}$$

und es ist $(0>)\ \varrho_1>\varrho_2>\varrho_3>\cdots$ das zugehörige Spektrum.

2° Nun liefert Lemma 1

$$[g(x)^2w(x)^{-1}\varphi''_\lambda(x)]'=\left\{\varrho_\lambda-\left[\frac{g'(x)}{w(x)}\right]'\right\}g(x)\varphi'_\lambda(x)\tag{5.5}$$
$$(a\leqq x\leqq b;\ \lambda=1,2,\ldots)$$

[45] Diese Annahme mag durch die Totalstetigkeit von $f(x)$ ersetzt werden.

mit (vgl. (5. 1), (5. 3))

$$(5.6)\qquad [g(x)\varphi'_\lambda(x)]'_{x=a} = g'(a)\varphi'_\lambda(a) + g(a)\varphi''_\lambda(a) = 0$$
$$(5.7)\qquad [g(x)\varphi'_\lambda(x)]'_{x=b} = g'(b)\varphi'_\lambda(b) + g(b)\varphi''_\lambda(b) = 0 \qquad (\lambda = 1, 2, \ldots);$$

man sieht zugleich, daß bei $\varrho \neq 0$ für jede Lösung $y = y(x)$ der Differentialgleichung

$$(5.8)\qquad [g(x)^2 w(x)^{-1} y']' = \left\{\varrho - \left[\frac{g'(x)}{w(x)}\right]'\right\} g(x) y \qquad (a \leqq x \leqq b),$$

welche den Randbedingungen

$$(5.9)\qquad [g(x)y]'_{x=a} = [g(x)y]'_{x=b} = 0$$

genügt, die primitive Funktion $Y = \int\limits_a^x y(t)dt$ (5. 4) befriedigt. Im Falle $\varrho = 0$ resultiert aber leicht $[g(x)y(x)]' \equiv 0$ $(a \leqq x \leqq b)$.

Daraus erhellt durch 1°, daß die Folge $g(x)^{-1}, \varphi'_1(x), \varphi'_2(x), \ldots$ als ein komplettes ($g$-vollständiges) Eigenfunktionensystem der $SL$-Aufgabe (5. 8), (5. 9) aufgefaßt werden kann.

3° Wir wenden auf (5. 5), (5. 6), (5. 7) die sogenannte *Liouvillesche Normaltransformation* an.[46] Man setze

$$(5.10)\qquad u = \int\limits_a^x \left(\frac{w(t)}{g(t)}\right)^{\frac{1}{2}} dt \quad (a \leqq x \leqq b), \qquad H = \int\limits_a^b \left(\frac{w(t)}{g(t)}\right)^{\frac{1}{2}} dt$$

und für die inverse Funktion

$$(5.11)\qquad x = V(u) \qquad (0 \leqq u \leqq H).$$

Schreibt man noch

$$(5.12)\qquad \varphi'_\lambda(x) = \sqrt{-\varrho_\lambda}\, g(x)^{-\frac{3}{4}} w(x)^{\frac{1}{4}} \varphi^*_\lambda(u)$$
$$(5.13)\qquad \varphi^*_\lambda(u) = (-\varrho_\lambda)^{-\frac{1}{2}} \left[\varphi'_\lambda(x) g(x)^{\frac{3}{4}} w(x)^{-\frac{1}{4}}\right]_{x=V(u)} \qquad (\lambda = 1, 2, \ldots),$$

ferner

$$(5.14)\qquad \varphi^*_0(u) = C_0[g(x)w(x)]^{-\frac{1}{4}}_{x=V(u)}, \quad C_0 = \left(\int\limits_a^b g(x)^{-1} dx\right)^{-\frac{1}{2}},$$

dann sind $y = \varphi^*_\lambda(u)$ $(\lambda = 0, 1, 2, \ldots)$ laut 2°, von konstanten Faktoren abgesehen, sämtliche Eigenfunktionen einer Differentialgleichung der Form

$$(5.15)\qquad \frac{d^2y}{du^2} + (\Xi(u) - \varrho)y = 0 \qquad (0 \leqq u \leqq H)$$

[46] Vgl. z. B. [19], S. 261—262.

mit

$$\Xi(u) = \left[\left(\frac{g'(x)}{w(x)}\right)' - \frac{1}{2}\left(\frac{g(x)}{w(x)}\right)' \frac{\Psi'(x)}{\Psi(x)} - \frac{g(x)}{w(x)} \frac{\Psi''(x)}{\Psi(x)}\right]_{x=V(u)}, \tag{5.16}$$

$$\Psi(x) = g(x)^{\frac{3}{4}} w(x)^{-\frac{1}{4}},$$

welche den (5.9) entsprechenden $SL$-Randbedingungen dritter Art genüge leisten. (Die Eigenwerte sind offenbar wieder $\varrho = 0, \varrho_1, \varrho_2, \ldots$.) Da wir ohne Mühe

$$\int_0^H \varphi_\lambda^*(u)^2 du = 1 \qquad (\lambda = 0, 1, 2, \ldots) \tag{5.17}$$

finden, so bilden $\varphi_\lambda^*(u)$ $(\lambda = 0, 1, 2, \ldots)$ ein vollständiges Orthonormalsystem in $(0, H)$.

4° Nach HOBSON gilt nun die Formel[47]

$$\varphi_\lambda^*(u) = \sqrt{\frac{2}{H}} \cos \frac{\pi\lambda}{H} u + \frac{\omega^*(u)}{\lambda} \sin \frac{\pi\lambda}{H} u + O\left(\frac{1}{\lambda^2}\right) \qquad (\lambda = 1, 2, \ldots), \tag{5.18}$$

wobei $\omega^*(u)$ eine von $\lambda$ unabhängige, in $0 \leqq u \leqq H$ stetige Funktion bedeutet und das letzte Glied ebenda gleichmäßig beschränkt ist.[48]

Daraus folgt vermöge des genannten Äquikonvergenztheorems von HAAR,[4] daß für eine *beliebige* $F(t) \in L(0, H)$ die Limesrelation

$$\lim_{n\to\infty} \left\{ \sum_{\lambda=0}^{n} \varphi_\lambda^*(u) \int_0^H F(t) \varphi_\lambda^*(t) dt - \right.$$

$$\left. - \left( \frac{1}{H} \int_0^H F(t)\, dt + \frac{2}{H} \sum_{\lambda=1}^{n} \cos \frac{\pi\lambda}{H} u \int_0^H F(t) \cos \frac{\pi\lambda}{H} t\, dt \right) \right\} = \tag{5.19}$$

$$= \lim_{n\to\infty} \int_0^H F(t) \left\{ \sum_{\lambda=0}^{n} \varphi_\lambda^*(t)\varphi_\lambda^*(u) - \frac{1}{H}\left(1 + 2\sum_{\lambda=1}^{n} \cos \frac{\pi\lambda}{H} t \cos \frac{\pi\lambda}{H} u\right) \right\} dt = 0$$

besteht, und sogar gleichmäßig für $0 \leqq u \leqq H$.

5° Es sei die Reihe

$$\sum_{\lambda=1}^{\infty} \left\{ \varphi_\lambda^*(t)\varphi_\lambda^*(u) - \frac{2}{H} \cos \frac{\pi\lambda}{H} t \cos \frac{\pi\lambda}{H} u - \right.$$

$$\left. - \sqrt{\frac{2}{H}} \frac{1}{\lambda} \left[ \omega^*(t) \sin \frac{\pi\lambda}{H} t \cos \frac{\pi\lambda}{H} u - \omega^*(u) \cos \frac{\pi\lambda}{H} t \sin \frac{\pi\lambda}{H} u \right] \right\} \tag{5.20}$$

$$(0 \leqq t \leqq H;\ 0 \leqq u \leqq H)$$

[47] Der genaue explizite Ausdruck von $\omega^*(u)$ ist leicht angebbar, aber für unsere Zwecke unnötig.

[48] Vgl. [17]; [19], S. 262.

[49] [13] M. I., S. 355.

betrachtet. Da ihr allgemeines Glied $O\left(\frac{1}{\lambda^2}\right)$ ist (vgl. (5. 18)), konvergiert sie nach WEIERSTRASS für die betreffenden $t$ und $u$ gleichmäßig und stellt eine stetige Funktion $M(t, u)$ dar. Somit gilt durch (5. 18)

$$(5.21)\qquad \sum_{\lambda=1}^{\infty}\left(\varphi_\lambda^*(0)\varphi_\lambda^*(t)-\frac{2}{H}\cos\frac{\pi\lambda}{H}t\right)=M(t,0)+\frac{\pi}{\sqrt{2H}}\,\omega^*(t)\left(1-\frac{t}{H}\right)$$
$$(0<t\leq H),$$

$$(5.22)\qquad \sum_{\lambda=0}^{\infty}\left(\varphi_\lambda(H)\varphi_\lambda^*(t)-\frac{2}{H}(-1)^\lambda\cos\frac{\pi\lambda}{H}t\right)=M(t,H)+\frac{\pi}{H\sqrt{2H}}\,t\,\omega^*(t)$$
$$(0\leq t<H).$$

Die Reihensummen linker Hand sind also in $(0, H)$ vorhanden, sogar stetig und in $[0, H]$ beschränkt; die Partialsummen beider Reihen bleiben in $[0, H]$, wie man auf Grund von $\left|\sum_{\lambda=1}^{n}\frac{\sin\lambda\xi}{\lambda}\right|\leq\int_0^{\pi}\frac{\sin t}{t}\,dt<2\;(-\infty<\xi<\infty,\, n=1,2,\ldots)$ leicht nachweist,[50] absolut unter einer von $n$ unabhängigen oberen Schranke.

Alles dies ermöglicht die Anwendung des Lebesgueschen Satzes über gliedweise Integration[51] im Falle von (5. 19); man erhält

$$\int_0^H F(t)\left\{\varphi_\lambda^*(0)\varphi_\lambda^*(t)-\frac{1}{H}+\sum_{\lambda=1}^{\infty}\left(\varphi_\lambda^*(0)\varphi_\lambda^*(t)-\frac{2}{H}\cos\frac{\pi\lambda}{H}t\right)\right\}dt=0,$$

$$\int_0^H F(t)\left\{\varphi_\lambda^*(H)\varphi_\lambda^*(t)-\frac{1}{H}+\sum_{\lambda=1}^{\infty}\left(\varphi_\lambda^*(H)\varphi_\lambda^*(t)-\frac{2}{H}(-1)^\lambda\cos\frac{\pi\lambda}{H}t\right)\right\}dt=0$$

und daraus, daß die Faktoren in Klammern neben $F(t)$ für $0<t<H$ identisch verschwinden müssen. Geht man endlich zu den arithmetischen Mitteln der Partialsummen über, so folgen wegen

$$(5.23)\qquad \lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}\left(1+2\sum_{\lambda=1}^{k}\cos\lambda\xi\right)=\lim_{n\to\infty}\frac{1}{n+1}\left(\frac{\sin\frac{1}{2}(n+1)\xi}{\sin\frac{1}{2}\xi}\right)^2=0$$
$$(0<\xi<2\pi)$$

die Relationen

$$(5.24)\qquad \lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}\sum_{\lambda=0}^{k}\varphi_\lambda^*(0)\varphi_\lambda^*(t)=\lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}\sum_{\lambda=0}^{k}\varphi_\lambda^*(H)\varphi_\lambda^*(t)=0$$
$$(0<t<H).$$

Sie gelten übrigens offenbar gleichmäßig in jedem Intervalle $[\varepsilon, H-\varepsilon]$ mit $\varepsilon>0$.

[50] Die oft benutzte gleichmäßige Beschränktheit der betreffenden Sinuspolynome kam bekanntlich zuerst bei FEJÉR zu ausgebreiteter Anwendung. (Vgl. [8]; ROGOSINSKI [27].)

[51] Vgl. z. B. [26], S. 37 ff.

6° Da, der Voraussetzung nach, $f(x) \in \mathrm{Lip}\, 1$ $(a \leqq x \leqq b)$, ist diese Funktion in $[a, b]$ stetig und *beschränkter Schwankung* und sogar, laut des wohlbekannten Lebesgueschen Satzes über monotone Funktionen,[52] ebenda *fast überall differenzierbar.*

Eine Teilintegration ergibt also (vgl. (5. 1), (5. 12), (5. 13))

$$
\begin{aligned}
\gamma_\lambda \sqrt{-\varrho_\lambda} = \sqrt{-\varrho_\lambda} \int_a^b f(t) w(t) \varphi_\lambda(t)\, dt = (-\varrho_\lambda)^{-\frac{1}{2}} \Big\{ \big[ f(t)\, g(t)\, \varphi'_\lambda(t) \big]_{t=a}^{t=b} + \\
+ \int_a^b f'(t) g(t) \varphi'_\lambda(t)\, dt \Big\} = f(a) \left( g(a) w(a) \right)^{\frac{1}{4}} \varphi^*_\lambda(0) - \\
- f(b) \left( g(b) w(b) \right)^{\frac{1}{4}} \varphi^*_\lambda(H) + \int_0^H \left[ f'(x) g(x)^{\frac{3}{4}} w(x)^{-\frac{1}{4}} \right]_{x = V(t)} \varphi^*_\lambda(t)\, dt
\end{aligned}
\tag{5. 25}
$$

$$(\lambda = 1, 2, \ldots),$$

wobei die Erklärung der Integrande passend ausgedehnt zu denken ist.[53] Es sei zunächst bemerkt, daß aus (5. 14), (5. 10)

$$
\begin{aligned}
f(a) \left( g(a) w(a) \right)^{\frac{1}{4}} \varphi^*_0(0) - f(b) \left( g(b) w(b) \right)^{\frac{1}{4}} \varphi^*_0(H) + \\
+ \int_0^H \left[ f'(x) g(x)^{\frac{3}{4}} w(x)^{-\frac{1}{4}} \right]_{x = V(t)} \varphi^*_0(t)\, dt = 0
\end{aligned}
\tag{5. 26}
$$

folgt.

Nun hat man formal

$$
\begin{aligned}
\sum_{\lambda=1}^{\infty} \gamma_\lambda \varphi'_\lambda(x) = g(x)^{-\frac{3}{4}} w(x)^{-\frac{1}{4}} \Big\{ f(a) \left( g(a) w(a) \right)^{\frac{1}{4}} \sum_{\lambda=0}^{\infty} \varphi^*_\lambda(0) \varphi^*_\lambda(u) - \\
- f(b) \left( g(b) w(b) \right)^{\frac{1}{4}} \sum_{\lambda=0}^{\infty} \varphi^*_\lambda(H) \varphi^*_\lambda(u) \Big\} + \\
+ g(x)^{-\frac{3}{4}} w(x)^{-\frac{1}{4}} \sum_{\lambda=0}^{\infty} \varphi^*_\lambda(u) \int_0^H \left[ f'(x) g(x)^{\frac{3}{4}} w(x)^{-\frac{1}{4}} \right]_{x = V(t)} \varphi^*_\lambda(t)\, dt.
\end{aligned}
\tag{5. 27}
$$

Das erste Glied rechts ist wegen (5. 24) $C_1$-summierbar zur $C_1$-Summe 0 für jedes $x \in (a, b)$, d. h. $u \in (0, H)$. Die Reihe im zweiten Gliede ist die Fouriersche Entwicklung der Funktion $[f'(x) g(x)^{\frac{3}{4}} w(x)^{-\frac{1}{4}}]_{x = V(u)}$ inbezug auf das Orthonormalsystem $\{\varphi^*_\lambda(u)\}$; da die trigonometrische Fourierreihe dieser Funktion nach der *Lebesgueschen Verallgemeinerung des Fejérschen Satzes*[54]

[52] Vgl. z. B. [26], S. 5—11.

[53] Etwa in der Weise, daß man den Punkten derjenigen Nullmenge, wo $f'(x)$ nicht existiert, den Funktionswert 0 zuordnet.

[54] Vgl. z. B. [16], S. 62.

fast überall, insbesondere an allen inneren Stetigkeitsstellen von $f'(x)$, zum Funktionswerte $C_1$-summierbar ist, ergibt das Haarsche Theorem (5. 19) dasselbe auf die fragliche Reihe.[55] Ähnlich schließt man auf gleichmäßige Summierbarkeit in (inneren) Stetigkeitsstrecken mit Hilfe von 5° bzw. des *Fejérschen Approximationssatzes.*

Somit sind die Behauptungen über $\Sigma \gamma_\lambda \varphi_\lambda^*(x)$ von (5. 27) ablesbar.

7° Um die Reihe $\Sigma \gamma_\lambda \varphi_\lambda^*(x)$ zu untersuchen, bemerken wir vor allem, daß die Benutzung von (5. 1), (5. 3) nebst der Transformation (5. 10), (5. 11) für $\varphi_\lambda(u) = [(g(x)w(x))^{\frac{1}{4}} \varphi_\lambda(x)]_{x=V(u)}$ die asymptotische Formel

$$(5.28) \quad \bar{\varphi}_\lambda(u) = \sqrt{\frac{2}{H}} \sin \frac{\pi\lambda}{H} u + \frac{\bar{\omega}(u)}{\lambda} \cos \frac{\pi\lambda}{H} u + O\left(\frac{1}{\lambda^2}\right) \qquad (\lambda = 1, 2, \ldots)$$

liefert, wobei $\omega(u)$ in $[0, H]$ stetig, von $\lambda$ frei und das Zeichen $O$ in $u$ gleichmäßig zu verstehen ist.[56]

Daraus schließt man einfach, daß unsere Reihe, welche einer Funktion $f(x) \in \mathrm{Lip}\, 1$ angehört, in $(a, b)$ konvergent ist und zwar in inneren Teilintervallen gleichmäßig konvergiert, ferner gleichmäßig $C_1$-summierbar ist. Was ihre Summe betrifft, so beachte man die nach (5. 19) für *jedes* in $[0, H]$ stetige $F(u)$ gültige Relation

$$(5.29) \qquad \lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} \sum_{\lambda=0}^{k} \varphi_\lambda^*(u) \int_0^H F(t)\varphi_\lambda^*(t)\,dt = F(u) \qquad (0 \leq u \leq H).$$

Multipliziert man mit $(g(x)w(x))^{\frac{1}{4}}$, differenziert[57] nach $x$ und schreibt $F(u) = \left[(g(x)w(x))^{-\frac{1}{4}} \int_a^x f(t)w(t)\,dt\right]_{x=V(u)}$, so folgt (vgl. (5. 1), (5. 13), (5. 14))

$$(5.30) \qquad \lim_{n\to\infty} \frac{1}{n+1} \sum_{k=1}^{n} \sum_{\lambda=1}^{k} \varphi_\lambda(x) \int_a^b f(t)w(t)\varphi_\lambda(t)\,dt = f(x) \qquad (a < x < b).$$

Wir haben also zugleich in $(a, b)$

$$(5.31) \qquad \sum_{\lambda=1}^{\infty} \gamma_\lambda \varphi_\lambda(x) = f(x),$$

w. z. b. w.

[55] Da $u = \int_a^x \left(\frac{w(t)}{g(t)}\right)^{\frac{1}{2}} dt$ und $x = V(u)$ in $[a, b]$ bzw. $[0, H]$ zur Klasse Lip 1 gehören, ersieht man leicht, daß Nullmengen auf der $x$- und $u$-Achse ineinander übergehen.

[56] Vgl. [19], S. 262.

[57] Dies ist bekanntlich gestattet, weil der Limes in (5. 30) nach den vorigen in jedem Inneren Teilintervalle von $(a, b)$ gleichmäßig existiert.

**5.2.** Wie man sieht, entspricht der soeben bewiesene Satz der Differentiation gewöhnlicher Fourierschen Sinusreihen. Es können offenbar ähnliche Sätze aufgestellt werden, in welchen ein anderes Paar homogen-linearer Randbedingungen, z. B. $\varphi_\lambda(a)=\varphi_\lambda(b)$, $\varphi_\lambda'(a)=\varphi_\lambda'(b)$ $(\lambda=1,2,\ldots)$ auftritt;[58] letzteres umfaßt auch den Fall der allgemeinen trigonometrischen Fourierreihe.

*(Eingegangen am 20. März 1954.)*

## Literaturverzeichnis


[1] J. Aczél, Eine Bemerkung über die Charakterisierung der klassischen orthogonalen Polynome, *Acta Math. Acad. Sci. Hung.*, **4** (1953), S. 315—321.

[2] S. Bochner, Über Sturm—Liouvillesche Polynomsysteme, *Math. Zeitschrift*, **29** (1929), S. 730—736.

[3] H. Cramér, *Mathematical methods of statistics* (Princeton, 1946).

[4] R. Courant und D. Hilbert, *Methoden der mathematischen Physik*, Bd. I (Berlin, 1924).

[5] J. Egerváry, Über die charakteristischen geometrischen Eigenschaften der Legendreschen und Tschebyscheffschen Polynome, *Archiv Math. u. Physik* (3), **27** (1918), S. 23—24.

[6] L. Fejér, Sur la différentiation de la série de Fourier, *Comptes Rendus Acad. Sci. Paris*, **134** (1902), S. 762—765.

[7] L. Fejér, Untersuchungen über Fouriersche Reihen, *Math. Annalen*, **58** (1904), S. 51—69.

[8] L. Fejér, Sur les singularités de la série de Fourier des fonctions continues, *Ann. École Normale Sup.*, **28** (1911), S. 64—103.

[9] Ph. Frank und R. Mises, *Die Differential- und Integralgleichungen der Mechanik und Physik*, Bd. I (Braunschweig, 1925).

[10] B. Gagaeff, Sur l'unicité de fonctions orthogonales invariant relativement à la dérivation, *Comptes Rendus Acad. Sci. Paris*, **188** (1929), S. 222—225.

[11] Б. Гагаев, О некоторых классах ортогональных функций, Известия Академии Наук СССР, Сер. матем., **10** (1946), S. 197—206.

[12] Б. Гнеденко, О единственности системы ортогональных функций, инвариантной относительно дифференцирования, Доклады Академии Наук СССР, **14** (1939), S. 159—161.

[13] A. Haar, Zur Theorie der orthogonalen Funktionensysteme (Habilitationsschrift), 1. und 2. Mitteil., *Math. Annalen*, **69** (1910), S. 331—371, bzw. ibid. **71** (1911), S. 38—53.

[14] W. Hahn, Über die Jacobischen Polynome und zwei verwandte Polynomklassen, *Math. Zeitschrift*, **39** (1935), S. 634—638.

[15] W. Hahn, Über höhere Ableitungen von Orthogonalpolynomen, ibid., **43** (1937), S. 101.

[16] G. H. Hardy and W. W. Rogosinski, *Fourier series* (Cambridge, 1944).

[17] E. W. Hobson, On a general convergence theorem and the theory of a function by series of normal functions, *Proc. London Math. Soc.* (2), **6** (1908) S. 349—395.


[58] Über die Asymptotik der Orthogonalfunktionen in den Fällen allgemeinerer Randbedingungen vgl. [31].

[18] D. Jackson, *Fourier-series and orthogonal polynomials* (Oberlin—Ohio, 1941).
[19] E. Kamke, *Differentialgleichungen, Lösungsmethoden und Lösungen*, Bd. I (Leipzig, 1951), 4. Auflage.
[20] E. Kamke, *Differentialgleichungen reeller Funktionen* (Leipzig, 1952), 2. Auflage.
[21] H. L. Krall, On derivatives of orthogonal polynomials, *Bull. Amer. Math. Soc.*, **42** (1936), S. 423—428.
[22] H. L. Krall, On higher derivatives of orthogonal polynomials, ibid., **42** (1936), S. 867—870.
[23] H. L. Krall, On higher derivatives of orthogonal polynomials. II, ibid., **47** (1941), S. 261—264.
[24] М. А. Лаврентьев и Л. А. Люстерник, Курс вариационного исчисления (Москва—Ленинград, 1950).
[25] D. C. Lewis, Orthogonal functions whose derivatives are also orthogonal, *Rend. Palermo* (2), **11** (1953), S. 159—168.
[26] F. Riesz et B. Sz.-Nagy, *Leçons d'analyse fonctionnelle* (Budapest, 1953), 2. Auflage.
[27] W. Rogosinsky, *Fouriersche Reihen* (Berlin—Leipzig, 1930), S. Göschen.
[28] J. Shohat, E. Hille and J. Walsh, *A bibliography on orthogonal polynomials*, Bull. Nat. Res. Council, **103** (Washington, 1940).
[29] G. Szegő, *Orthogonal polynomials* (New York, 1939).
[30] M. S. Webster, Orthogonal polynomials with orthogonal derivatives, *Bull. Amer. Math. Soc.*, **44** (1938), S. 880—888.
[31] A. C. Zaanen, On some orthogonal systems of functions, *Compositio Math.*, **7** (1940), S. 253—282.
[32] A. Zygmund, *Trigonometrical series* (Warszawa—Lwów, 1935).

## О НЕКОТОРЫХ СВОЙСТВАХ ОРТОГОНАЛЬНЫХ СИСТЕМ И О СОБСТВЕННЫХ ЗНАЧЕНИЯХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ ТИПА ШТУРМА—ЛИУВИЛЛЯ

М. Миколаш (Будапешт)

(Резюме)

С точки зрения применений не лишен интереса факт, что тригонометрическая ортогональная система и системы „классических“ ортогональных полиномов при дифференцировании и при интегрировании (выполненном после умножения на весовую функцию) переходят опять в ортогональные системы, ведь это является причиной того, что ортогональные ряды по элементам указанных систем, в аналогии с степенными рядами, сохраняют свой „характер“ при почленном дифференцировании и интегрировании.

Настоящая работа посвящена изучении тех ортогональных системах пространства $L^2$, которые обладают указанными свойствами важнейших ортогональных систем. Более подробно исследуются следующие проблемы:

I. Пусть $\varphi_\lambda(x) \in L^2_w(a, b)$ $(\lambda = 1, 2, \ldots)$ — ортогональная система, а $w(x)$ — неотрицательная весовая функция. Пусть даны некоторые постоянные $c_1, c_2, \ldots$ и другая весовая функция $Q(x) \geqq O$ $(a \leqq x \leqq b)$. Спрашивается необходимое и достаточное условие для того, чтобы — при обозначении $\Phi_\lambda(x) = \int\limits_a^x w(t)\,\varphi_\lambda(t)\,dt$ $(\lambda = 1, 2, \ldots)$ — система $\{\Phi_\lambda(x) + c_\lambda\}$ или $1, \{\Phi_\lambda(x) + c_\lambda\}$ также была ортогональной системой на $(a, b)$ относительно $Q(x)$.

II. Пусть $\varphi_\lambda(x)$ ($\lambda = 1, 2, \ldots$) система функций дифференцируемых на $(a, b)$, ортогональная на этом интервале относительно весовой функции $w(x)$ ($w$-ортогональная). Зададимся также другой весовой функцией $g(x) \geqq 0$ ($a \leqq x \leqq b$). Спрашивается условие необходимое и достаточное для того, чтобы система $\{\varphi_\lambda'(x)\}$ была $g$-ортогональна на $(a, b)$.

III. Требуется задать и изучать по возможности широкий класс ортогональных систем, элементы которого обладают свойствами I и II.

IV. На основе изложенного следует охарактеризовать „классические" ортогональные системы, в частности многочлены Якоби, Лагерра и Эрмита, а также собственные функции Штурма—Лиувилля.

V. Что можно сказать о сходимости или о суммируемости (общего) ряда Фурье и его производного ряда, если известно, что ряд полученный почленным дифференцированием также ортогонален?

Эти вопросы кажутся новыми, однако следует упоминать о многочисленных частных результатах, дающих различные от общепринятых определения классически ортогональных полиномов (см. [28]). Сюда относится и одно замечание Гагаева о системе 1, $\cos x$, $\sin x$, $\cos 2x$, $\sin 2x$, .... Исследования, отчасти примыкающие к II, можно найти и в новейшей работе Д. С. Льюиса [25], о которой автор узнал при корректуре.

Главные результаты настоящей работы состоят в следующем:

1. Полная ортогональная система обладает свойствами I и II тогда и только тогда, если ее элементы удовлетворяют почти всюду некоторому интегральному уравнению, или (что в случае непрерывных функций этому эквивалентно) дифференциальному уравнению типа Штурма—Лиувилля

$$[g(x)y']' = \varrho_\lambda w(x) y,$$

связанному удобными краевыми условиями. (См. теоремы 1, 2 и 3.)

2. Таким образом легко найти некоторый общий класс ортогональных систем, из которых дифференцированием, также как и интегрированием, получается опять ортогональная система. Итак, класс этих „обыкновенных систем $SL$" содержит „классические" ортогональные системы. (Теоремы 4 и 5.)

3. В качестве применения автору удается дать общую характеристику тригонометрической системы, собственных функций Штурма—Лиувилля первого и второго рядов и классических ортогональных полиномов, использующую только свойства общей природы (ограниченность, непрерывность, дифференцируемость, ортогональность, полнота) без требования выполненности дифференциального уравнения или другого специального соотношения. (Теоремы 6, 7, 8.) Пример: С точностью до перестановок и до тривиальных линейных преобразований, система 1, $\cos x$, $\sin x$, $\cos 2x$, $\sin 2x$, ... является единственной полной ортогональной системой в интервале $(0, 2\pi)$, для которой система первых и вторых производных, дополненная единицей, также полная ортогональная система на том же интервале.

4. Наконец, в связи с проблемой V, доказывается теорема, которая может рассматриваться как обобщение результата, полученного Фейером для обыкновенных рядов Фурье ([6], [7]). В этой теореме утверждается, по существу, следующее: если для некоторой функции $f(x) \in \mathrm{Lip}\,1$ ($a \leqq x \leqq b$) ее ряд Фурье по системе $\{\varphi_\lambda(x)\}$ ($\varphi_\lambda(a) = \varphi_\lambda(b) = 0$, $\lambda = 1, 2, \ldots$) $w$-ортогональной и $w$-полной на $(a, b)$ переходит при почленном дифференцировании в ряд также ортогональный на $(a, b)$, то ряд Фурье сходится на $(a, b)$ и его сумма равна $f(x)$, производный же ряд почти всюду является суммируемым $C_1$ и его $C_1$-сумма равна $f'(x)$; суммируемость является равномерной во всяком замкнутом интервале непрерывности $f'(x)$.

# ÜBER TRIGONOMETRISCHE INTEGRALE, WELCHE GANZE FUNKTIONEN MIT NUR REELLEN NULLSTELLEN DARSTELLEN

Von

LJUBOMIR ILIEFF (Sofia)

(*Vorgelegt von* P. TURÁN)

**1.** Es sei $E(a)$ die Gesamtheit der Funktionen $f(t)$, die im Intervall $(0, a)$, $a>0$, nicht negativ, R-integrierbar sind, und der Bedingung genügen, daß die ganze Funktion

$$F(z)=\int_0^a f(t)\cos tz\,dt \tag{F}$$

nur reelle Nullstellen besitzt.

Es sei $x=x(t)$ ein reelles, gerades Polynom von $t$ oder auch eine reelle, gerade, ganze Funktion, so daß

a) $x(a)=0$,

b) $x'(it)$ ein Polynom mit nur reellen Nullstellen oder eine ganze Funktion, die eine Grenze solcher Polynome ist.

Da $x'(t)$ eine reelle Funktion ist, die außer $t=0$ keine reellen Nullstellen besitzt, so folgt, daß sie in $-\infty<t<0$ und in $0<t<+\infty$ ihr Vorzeichen nicht wechselt. $x(t)$ genügt also auch der Bedingung:

c) In jedem der Intervalle $(-\infty, 0)$, $(0, +\infty)$ ist die Funktion $x(t)$ entweder streng wachsend oder streng abnehmend.

Da, wegen der Bedingung b), $x'(a)\neq 0$ ist, so folgt aus c), daß $x(t)$ auch der Bedingung genügt:

d) $x(t)$ besitzt nur zwei reelle Nullstellen: $t_1=a$ und $t_2=-a$. Diese Nullstellen sind einfach.

Es sei $A(a)$ die Klasse der reellen, geraden Funktionen $x(t)$, die den Bedingungen a) und b) genügen.

Gemäß der Bedingung c) hat jede Funktion von der Klasse $A(a)$ im Intervall $[0, a]$ eine inverse Funktion $t=t(x)$.

Es sei $f_0\in E(a)$, $x(t)\in A(a)$ und $t=t(x)$ sei die inverse Funktion von $x(t)$ im Intervall $[0, a]$.

Man setze

$$\varphi_0(x) = f_0(t), \quad f_1(t) = \varphi_1(x) = \int_0^x \varphi_0(v)\,dv$$

und bezeichne, der Kürze wegen,

$$f_1(t) = L(f_0, x).$$

SATZ I. *Wenn* $f_0(t) \in E(a)$, $x(t) \in A(a)$ *und* $f_1(t) = L(f_0, x)$, *so haben wir* $f_1(t) \in E(a)$.

BEWEIS. Da $f_0(t) \in E(a)$, so besitzt die ganze Funktion

$$F_0(z) = \int_0^a f_0(t) \cos tz\,dt$$

nur reelle Nullstellen. Man setze $f_0(-t) = f_0(t)$ für $0 \leqq t \leqq a$. Dann ist

$$F_0(z) = \frac{1}{2} \int_{-a}^{a} f_0(t) e^{itz}\,dt.$$

Gemäß der Bedingung b) und einem Ergebnis von G. PÓLYA [1] hat auch die ganze Funktion

$$\int_{-a}^{a} f_0(t)\, x'(t) e^{itz}\,dt = 2i \int_0^a f_0(t) x'(t) \sin tz\,dt$$

auter reelle Nullstellen.

Es ist

$$\int_0^a f_0(t) x'(t) \sin tz\,dt = \int_{t=0}^{t=a} \varphi_0(x) \sin tz\,dx = \int_{t=0}^{t=a} \sin tz\,d\varphi_1(x) =$$

$$= \varphi_1(x) \sin tz \Big|_{t=0}^{t=a} - z \int_0^a \varphi_1(x) \cos tz\,dt = -z \int_0^a f_1(t) \cos tz\,dt,$$

da $\varphi_1[x(a)] = \varphi_1(0) = 0$. Der Satz ist bewiesen.

2. Mit Hilfe des Satzes I bekommt man also aus jeder Funktion von der Form (F) mit nur reellen Nullstellen eine neue Funktion derselben Art.

Die Funktion $x(t) \in A(a)$ genüge noch der Bedingung: a′) $x(0) > 0$. In diesem Falle nimmt die Funktion $x(t)$ im Intervall $[0, a]$ zu, so daß $x^{-\alpha}$, wo $0 \leqq \alpha < 1$ ist, eine wachsende Funktion in demselben Intervall darstellt. Nach wohlbekannten Ergebnissen von G. PÓLYA [2] folgt, daß die Funktion $f_0(t) = x^{-\alpha}$ zu der Klasse $E(a)$ gehört. Aus dem Satze I folgt in diesem Falle, daß die Funktionen $f_0(t) = \varphi_0(x) = x^{-\alpha}$, $(1-\alpha) f_1(t) = (1-\alpha)\varphi_1(x) = x^{1-\alpha}, \ldots$, $(1-\alpha)(2-\alpha)\ldots(n-\alpha) f_n(t) = x^{n-\alpha}, \ldots$ zu der Klasse $E(a)$ gehören. Da zu jedem $\lambda > -1$ immer ein ganzes $n$ und eine Zahl $\alpha$ mit $0 \leqq \alpha < 1$ gehört, so daß $\lambda = n - \alpha$ ist, bekommt man die

Folgerung. *Wenn $x(t) \in A(a)$, $x(0) > 0$ und $\lambda > -1$ besteht, so hat die ganze Funktion*

$$\int_0^a x^\lambda \cos tz\,dt$$

*nur reelle Nullstellen.*

So gehört z. B. die Funktion $x = 1 - t^{2q}$, wo $q$ eine positive ganze Zahl bezeichnet, zu der Klasse $A(1)$, wobei $x(0) > 0$. Aus der Folgerung bekommt man das bekannte Ergebnis von G. Pólya [1], nach dem die ganze Funktion

$$\int_0^1 (1 - t^{2q})^\lambda \cos tz\,dt,$$

falls $\lambda > -1$ und $q$ eine ganze positive Zahl bezeichnet, lauter reelle Nullstellen besitzt.

**3.** Es sei $f(t)$ ein reelles, gerades Polynom, so daß $f'(it)$ lauter reelle Nullstellen besitzt. In diesem Falle hat $f'(t)$ im Intervall $(0, +\infty)$ ein und dasselbe Vorzeichen, so daß $f(t)$ daselbst streng monoton ist. Da aber auch $f''(t) \neq 0$ für $0 < t < +\infty$ ist, haben wir entweder $\lim\limits_{t \to \infty} f(t) = +\infty$ oder $\lim\limits_{t \to \infty} f(t) = -\infty$. Im ersten Falle, wenn man eine passende Konstante zufügt, kann man annehmen, daß $f(t) \geqq 0$; und im zweiten Falle, daß $f(t) \leqq 0$. Es sei $f(t) \geqq 0$. Dann stellt $f(t)$ eine wachsende Funktion im Intervall $(0, +\infty)$ dar. Ist also $n > f(0)$ eine beliebige Zahl, so existiert wegen $\lim\limits_{t \to \infty} f(t) = +\infty$ eine Zahl $a_n > 0$, so daß $f(a_n) = n$. Wegen der Monotonie von $f(t)$ gibt es nur ein einziges $a_n$, wo $\lim\limits_{n \to +\infty} a_n = +\infty$. Es ist also $1 - \dfrac{f(t)}{n} \in A(a_n)$. Jetzt liefert unsere Folgerung das Ergebnis, daß die ganze Funktion

$$\int_0^{a_n} \left(1 - \frac{f(t)}{n}\right)^n \cos tz\,dt$$

nur reelle Nullstellen besitzt. Da im Intervall $(0, a_n)$ die Ungleichungen $0 < f(t) < n$, $0 < \left(1 - \dfrac{f(t)}{n}\right)^n < e^{-f(t)}$ bestehen, so bekommt man, daß die ganze Funktion $\displaystyle\int_0^\infty e^{-f(t)} \cos tz\,dt$ lauter reelle Nullstellen besitzt. Durch Grenzübergang bekommt man:

Satz II. *Wenn $f(t)$ eine nichtnegative, gerade Funktion bezeichnet, so daß $f'(it)$ ein Polynom mit nur reellen Nullstellen oder eine Grenze solcher Polynome ist, so besitzt die ganze Funktion*

$$\int_0^\infty e^{-f(t)} \cos tz\,dt$$

*nur reelle Nullstellen.*

Die Funktion $f(t) = a \operatorname{ch} t = \frac{a}{2}(e^t + e^{-t})$, $a > 0$, genügt den Bedingungen von Satz II. So bekommt man aus demselben Satz das bekannte Ergebnis von G. PÓLYA [3], nach dem die Funktion

$$\int_0^\infty e^{-a \operatorname{ch} t} \cos tz\,dt$$

nur reelle Nullstellen besitzt.

*(Eingegangen am 12. November 1954.)*

## Literaturverzeichnis

[1] G. PÓLYA, Über trigonometrische Integrale mit nur reellen Nullstellen, *Journal f. reine u. angew. Math.*, **158** (1927), S. 6—18.
[2] G. PÓLYA, Über die Nullstellen gewisser ganzer Funktionen, *Math. Zeitschr.*, **2** (1918), S. 352—383.
[3] G. PÓLYA, Bemerkung über die Integraldarstellung der Riemannschen $\xi$-Funktion, *Acta Math.*, **48** (1926), S. 305—317.

## О ТРИГОНОМЕТРИЧЕСКИХ ИНТЕГРАЛАХ, ПРЕДСТАВЛЯЮЩИХ ЦЕЛЫЕ ФУНКЦИИ, ИМЕЮЩИЕ ТОЛЬКО ВЕЩЕСТВЕННЫХ КОРНЕЙ

Л. Илиев (София)

(Резюме)

В настоящей работе дается простое доказательство двух теорем.

Теорема I. Если $x(t)$ — четная целая функция, для некоторого положительного $a$ $x(a) = 0$, а $x'(it)$ является пределом многочленов, все корни которых вещественны, то из того, что все корни интеграла

$$\int_0^a f_0(t) \cos tz\,dt$$

вещественны, следует, что

$$\int_0^a f_1(t) \cos tz\,dt$$

также обладает этим свойством, где

$$\varphi_0(x) = f_0(t), \qquad f_1(t) = \int_0^x \varphi_0(v)\,dv.$$

Теорема II. Если $f(t) \geqq 0$ — четная функция и $f'(it)$ является пределом многочленов, имеющих только вещественные корни, то все корни интеграла

$$\int_0^\infty e^{-f(t)} \cos tz\,dt$$

вещественны.

# PROJECTIONS OF PROBABILITY DISTRIBUTIONS

By

WALTER M. GILBERT (Pullman, USA)

(*Presented by* A. RÉNYI)

## § 1. Introduction

A two dimensional probability distribution generates a family of one dimensional distributions by (orthogonal) projection upon straight lines in the plane of the distribution. Since projections on parallel lines are the same it is sufficient to consider only lines through the origin. It is of interest to see what properties of the "parent" distribution can be determined by an examination of some or all of the projection distributions. If all of the projections are known the parent is completely determined, by a theorem of H. CRAMÉR and H. WOLD [1]. A. RÉNYI [3] has given a condition under which a two dimensional distribution is determined by any infinite set of projections. His condition is that the two dimensional characteristic function be an analytic function of $\theta$ for every $r$, where $r$ and $\theta$ are polar coordinates. In this note we show that in general a distribution need not be determined by an infinite set of projections, and we give another condition which insures determination. We also show that if two projections of a distribution have analytic characteristic functions, then so have all the others, and the two dimensional characteristic function is itself analytic.

## § 2. Determination of the parent

The fundamental theorem on the subject is due to CRAMÉR and WOLD [1]. It states that any distribution is determined by the set of all of its projections. This is established by showing that the two dimensional characteristic function, considered on the line through the origin of inclination $\theta$, coincides with the characteristic function of the projection on a line of inclination $\theta$. RÉNYI's problem is thus equivalent to finding conditions under which a two dimensional characteristic function is determined by its values on an infinite set of lines through the origin.

That an infinite number of projections need not determine a two dimensional distribution, is clear from a slight modification of a well-known example of two distinct one dimensional characteristic functions which coincide on a finite interval. Such an example, attributed to KHINTCHINE, is given in LÉVY [2]. Let $\varphi_1(t)$ equal $1-|t|$ for $|t| \leqq 1$ and zero for $|t| \geqq 1$, and let $\varphi_2(t)$ coincide with $\varphi_1(t)$ for $|t| \leqq 1$ but be periodic with period 2. (The distributions of which these are the characteristic functions are given in [2].) To form our example we consider the two dimensional characteristic functions $\psi_1(t_1, t_2) = \varphi_1(t_1)\varphi_1(t_2)$ and $\psi_2(t_1, t_2) = \varphi_1(t_1)\varphi_2(t_2)$. Then $\psi_1$ and $\psi_2$ coincide along all lines $L_\theta$ through the origin for which $\pi/4 \leqq \theta \leqq 3\pi/4$, but differ along all other lines. It may be noted that neither of the distributions associated with $\psi_1$ and $\psi_2$ has any moments.

On the other hand, we may prove the following theorem.

THEOREM 1. *If the moments of a distribution $F(x_1, x_2)$ are those of a determined moment problem, then $F(x_1, x_2)$ is determined by its projections on any infinite set of distinct lines through the origin.*

PROOF. It is sufficient to show that the values at the origin of all of the derivatives of the characteristic function $\psi(t_1, t_2)$ of $F(x_1, x_2)$ may be calculated from the characteristic functions $\varphi_i(t)$ of a denumerable set of the projection distributions $F_i(x)$.

Consider then characteristic functions $\varphi_i(t)$ $(i = 1, 2, \ldots, n+1)$ of $n+1$ projections on $n+1$ distinct lines, where the line $L_i$ makes an angle $0 \leqq \theta_i < \pi$ with the $x_1$ axis. By the theorem of CRAMÉR and WOLD, $\varphi_i(t)$ is the value of $\psi(t_1, t_2)$ along the line through the origin making an angle $\theta_i$ with the $t_1$ axis. We may differentiate each of the $\varphi_i(t)$ $n$ times and put $t = 0$. This yields $n+1$ linear equations:

$$\varphi_i^{(n)}(0) = \sum_{j=0}^{n} C_j^n \cos^{n-j}\theta_i \sin^j \theta_i \,.\, \psi^{(n-j,\,j)}(0, 0).$$

These may be solved for $\psi^{(n-j,j)}(0, 0)$ provided the determinant of the coefficients is not zero. To show that the determinant is not zero, suppose for the moment that no $\theta_i$ is $\pi/2$. We may then divide the $j+1$-st column by $C_j^n$ and the $i$-th row by $\cos^n \theta_j$. Letting $y_i = \operatorname{tg}\theta_i$ we have the Vandermonde determinant, which will be zero only for $\operatorname{tg}\theta_i = \operatorname{tg}\theta_j$ with $i \neq j$. But since $0 \leqq \theta_i < \pi$ this will occur only for $\theta_i = \theta_j$, and we have supposed the $\theta_i$'s distinct. If some $\theta_i = \pi/2$, all the elements of the $i$-th row are zero save the $n+1$-st, which is equal to one. We may then apply the above argument to the cofactor of this element.

A distribution may satisfy RÉNYI's condition without satisfying ours. For example, the characteristic function associated with the distribution with density $\dfrac{1}{\pi(1+r^2)^2}$ is constant for fixed $r$, but the distribution has no proper moments. An example of a distribution whose moment problem is determined

but whose characteristic function is not analytic (if such exists) would serve to show the converse.

A condition which insures that a two dimensional moment problem is determined has been given by CRAMÉR and WOLD [1], p. 291. We may apply it to obtain a corollary which imposes requirements on the projections rather than the two dimensional distribution.

COROLLARY. Suppose given the projections of a two dimensional distribution on a countable set of distinct lines through the origin. Let $\mu_n^i$ be the $n$-th moment of the $i$-th projection. Suppose that the $j$-th and $k$-th projections are such that

$$\sum_{n=1}^{\infty} (\mu_{2n}^j + \mu_{2n}^k)^{-\frac{1}{2n}} = \infty .$$

Then the two dimensional distribution is uniquely determined by the set of projections.

## § 3. Projections with analytic characteristic functions

The proof of the CRAMÉR—WOLD theorem suggests the following theorem, which, however, does not lead to unique determination of the original distribution.

THEOREM 2. *If two projections of a two dimensional distribution have analytic characteristic functions, then the same is true for any projection, and the characteristic function of the parent distribution is an analytic function of two complex variables in a neighbourhood of the origin.*

PROOF. In WINTNER [4] it is shown that a characteristic function is analytic for $t < q$ if and only if $\mu_{2n} = O[(2n)!p^{-2n}]$ as $n \to \infty$ for every $0 < p < q$. Following [4] and using Stirling's formula, $(\mu_{2n})^{\frac{1}{2n}} \leqq kn$ is equivalent to $\mu_{2n} = O[(2n)!(ke)^{2n}2^{-2n}]$. Thus $(\mu_{2n})^{\frac{1}{2n}} \leqq kn$ is necessary and sufficient for the analyticity of the characteristic function, and the $k$ determines a minimum radius of convergence of the Taylor series about the origin.

It is no restriction to suppose that the two lines on which $\psi(t_1, t_2)$ is analytic are the coordinate axes, since a linear transformation will make this so. We have then $(\mu_{2n}^1)^{\frac{1}{2n}} \leqq k_1 n$ and $(\mu_{2n}^2)^{\frac{1}{2n}} \leqq k_2 n$. Let $\mu_{2n}^1 + \mu_{2n}^2 = \lambda_{2n}$. Then $\lambda_{2n} < [(k_1)^{2n} + (k_2)^{2n}]\, n^{2n} \leqq (k_1 + k_2)^{2n} n^{2n}$ so that $\lambda_{2n}^{\frac{1}{2n}} < kn$ also. Let $F_c$ be the distribution resulting from a projection on a line with direction cosines $c_1$ and $c_2$. Then

$$\mu_{2n}(F_c) = \iint (c_1 x_1 + c_2 x_2)^{2n}\, dF(x_1, x_2) \leqq 2^{2n} \iint (x_1^{2n} + x_2^{2n})\, dF(x_1, x_2) = 2^{2n} \lambda_{2n}$$

and so there is a $K$ such that for any projection $[\mu_{2n}(F_\varphi)]^{\frac{1}{2n}} < Kn$, and a corresponding uniform minimum radius of convergence. It follows that the multiple Taylor series converges for any point within a circle about the origin in the real $(t_1, t_2)$ plane and therefore within a polycylinder in the space of two complex variables.

The argument of the last paragraph is easily applied to show also that if any two projections have $n$ moments, then so has any other projection and so has the parent distribution.

STATE COLLEGE OF WASHINGTON

(*Received 12 November 1954*)

## References

[1] H. Cramér and H. Wold, Some theorems on distribution functions, *Journal of the London Mathematical Society*, **11** (1936), pp. 290—294.
[2] P. Lévy, *Théorie de l'addition des variables aléatoires* (Paris, 1937).
[3] A. Rényi, On projections of probability distributions, *Acta Math. Acad. Sci. Hung.*, **3** (1952), pp. 131—142.
[4] A. Wintner, *The Fourier transforms of probability distributions* (Baltimore, 1947).

## О ПРОЕКЦИЯХ РАСПРЕДЕЛЕНИЙ ВЕРОЯТНОСТЕЙ

В. М. Джилберт (Пульман, США)

(Резюме)

А. Реньи в работе [3] ставил вопрос об условиях, при которых двумерное распределение вероятностей определяется его проекциями на бесконечное множество различных прямых, проходящих через начало координат; он же дал достаточное условие этого. В настоящей работе приводится, с одной стороны, простой пример распределения, не являющегося определенным бесконечным множеством его различных проекций, а с другой стороны доказывается, что если для двумерного распределения (двумерная) проблема моментов является определенным, то распределение определяется любым бесконечным множеством его различных проекций. Доказывается, наконец, что если характеристические функции двух проекций двумерного распределения аналитичны на интервале содержащем точку $O$, то этим же свойством обладают все проекции данного распределения, а двумерная характеристическая функция распределения также аналитична в области содержащей начало координат.

# REFINEMENT OF A PROBABILITY LIMIT THEOREM AND ITS APPLICATION TO BESSEL FUNCTIONS

By

M. FISZ (Warszawa)

(*Presented by* A. RÉNYI)

**1.** We consider two independent Poisson random variables $X_1$ and $X_2$ distributed according to the law

(1) $$P(X_i=k)=e^{-w_i}\frac{w_i^k}{k!},$$

where $w_i>0$; $i=1,2$; $k=0,1,2,\dots$.

The difference $X_2-X_1$ is, as $w_1\to\infty$ and $w_2\to\infty$, asymptotically normal. This has been shown by J. O. IRWIN [1] for the special case $w_1=w_2$, and for the general case by M. FISZ [2] and [3].

The problem of the refinement of these limit theorems is dealt with in this paper. The results obtained are given in formulae (7) and (11) and as a byproduct the asymptotical formula (13) for Bessel functions with purely imaginary argument has been obtained.

**2.** Let us consider the Poisson random variables mentioned above and let us suppose that

(2) $$w_i=nd_i \qquad (i=1,2),$$

where $d_1$ and $d_2$ are some real and positive constants. Let us denote

(3) $$z=\frac{k-(w_2-w_1)}{\sqrt{w_1+w_2}},$$

(4) $$\varphi(z)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2}\right),$$

(5) $$\mu_{r+2}=\begin{cases}(w_1+w_2)^{-\frac{r}{2}} & (r=2,4,6,\dots),\\ (w_2-w_1)(w_1+w_2)^{-\frac{r+2}{2}} & (r=1,3,5,\dots).\end{cases}$$

Finally let $F_r[-\varphi(z)]$ denote the function which can be obtained for each value of $r=1,2,\ldots$ from the identity:

$$\exp\left\{\sum_{r=1}^{\infty}\frac{\mu_{r+2}}{(r+2)!}\left[-\varphi(z)\right]^{(r+2)}\left(\frac{1}{\sqrt{n}}\right)^r\right\}=1+\sum_{r=1}^{\infty}F_r[-\varphi(z)]\left(\frac{1}{\sqrt{n}}\right)^r \tag{6}$$

by comparing the coefficients of $\frac{1}{\sqrt{n}}$ on both sides of this identity. In the formula (6) we have

$$[-\varphi(z)]^{(r+2)}=(-1)^{r+1}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2}\right)H_{r+2}(z),$$

where $H_{r+2}(z)$ is the Hermite polynomial of order $r+2$.

The following theorem will be proved.

THEOREM. *Let $X_1$ and $X_2$ be two independent random variables distributed according to* (1) *and let the equalities* (2) *hold. Then, as $w_i\to\infty$ $(i=1,2)$, the asymptotical formula*

$$\begin{aligned} P_{w_1w_2}(k)&=P(X_2-X_1=k)=\\ &=\frac{1}{\sqrt{w_1+w_2}}\left\{\varphi(z)+\sum_{r=1}^{u}F_r[-\varphi(z)]\right\}+o\left[\frac{1}{(w_1+w_2)^{\frac{u+1}{2}}}\right] \end{aligned} \tag{7}$$

*holds for each integer $u\geqq 1$.*

PROOF. We deduce the assertion of our Theorem from the following theorem given by C. G. ESSEEN [4]:

Let $\xi_1,\xi_2,\ldots$ be independent and equally distributed, integer-valued random variables having the mean value $m$ and the standard deviation $\sigma\neq 0$ and finite moments of order $v\geqq 3$ and let the greatest common divisor of the differences of pairs of all possible values of $\xi$ be equal to 1. Then the asymptotical formula

$$\begin{aligned} &P(\xi_1+\xi_2+\cdots+\xi_n=k)=\\ &=\frac{1}{\sigma\sqrt{n}}\left\{\varphi(z)+\sum_{r=1}^{v-2}\frac{1}{n^{\frac{r}{2}}}U_r[-\varphi(z)]\right\}+o\left(\frac{1}{n^{\frac{v-1}{2}}}\right) \end{aligned} \tag{8}$$

holds, where

$$z=\frac{k-mn}{\sigma\sqrt{n}},$$

$\varphi(z)$ is given by (4) and the functions $U_r$ can be determined for each $r$ from the formula

$$\exp\left\{\sum_{r=1}^{\infty}\frac{\lambda_{r+2}}{(r+2)!}[-\varphi(z)]^{(r+2)}\left(\frac{1}{\sqrt{n}}\right)^r\right\}=1+\sum_{r=1}^{\infty}U_r[-\varphi(z)]\left(\frac{1}{\sqrt{n}}\right)^r, \tag{9}$$

in which $\lambda_{r+1}$ is the $(r+2)$-th semiinvariant of the variable $\frac{\xi}{\sigma}$.

The applicability of the theorem of ESSEEN to our Theorem follows from the fact that the addition theorem is satisfied for Poisson variables and their differences. Taking into account the relations (2), we can write

$$X_i = \sum_{k=1}^{n} \eta_{ik} \qquad (i = 1, 2),$$

where $\eta_{ik}$ $(k = 1, 2, \ldots)$ are independent and equally distributed Poisson variables having the mean value $d_i$ $(i = 1, 2)$. The variables

$$\xi_k = \eta_{2k} - \eta_{1k} \qquad (k = 1, 2, \ldots)$$

satisfy all the assumptions of the theorem of ESSEEN and have finite moments of arbitrary order. The semiinvariants of $\xi_k$ are given by the formulae

$$\lambda_{r+2} = \begin{cases} n^{\frac{r}{2}} (w_1 + w_2)^{-\frac{r}{2}} & (r = 2, 4, \ldots), \\ n^{\frac{r}{2}} (w_2 - w_1)(w_1 + w_2)^{-\frac{r+2}{2}} & (r = 1, 3, \ldots). \end{cases} \tag{10}$$

Taking into account the formula (10), we obtain from (8) and (9) the formula (7).

Our Theorem is thus proved.

In the case when $X_1$ and $X_2$ are equally distributed, thus when $w_1 = w_2$, all coefficients $\mu_{r+2}$ given by (5) vanish for odd values of $r$. It is then convenient to write the formula (7) in the following form:

$$P_w(k) = P(X_2 - X_1 = k) = \frac{1}{\sqrt{2w}} \left\{ \varphi(z) + \sum_{r=1}^{u} F_{2r}[-\varphi(z)] \right\} + o\left(\frac{1}{w^{u+1}}\right). \tag{11}$$

Let us observe that the exact value of $P_w(k)$ is given by the formula

$$P_w(k) = e^{-2w} \sum_{s=0}^{\infty} \frac{w^{2s+k}}{s!\,(s+k)!} = e^{-2w} I_k(2w), \tag{12}$$

where $I_k(2w)$ is a Bessel function with purely imaginary argument. From (11) and (12) we obtain the following asymptotical formula:

$$I_k(2w) = e^{2w} \left[ \frac{1}{\sqrt{2w}} \left\{ \varphi(z) + \sum_{r=1}^{u} F_{2r}[-\varphi(z)] \right\} + o\left(\frac{1}{w^{u+1}}\right) \right]. \tag{13}$$

Let us observe that the formula (13) should be used for large $k$ and large $w$, namely when $k$ is of order $\sqrt{2w}$.

We shall finally give two short expansions of the formulae (7) and (12). We find namely

$$F_1 = \frac{\mu_3}{3!} [-\varphi(z)]^{(3)} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}} \frac{w_2 - w_1}{(w_1 + w_2)^{\frac{3}{2}}} H_3(z),$$

$$F_2 = \frac{\mu_4}{4!} [-\varphi(z)]^{(4)} + \frac{1}{2!} \frac{\mu_3^2}{(3!)^2} [-\varphi(z)]^{(6)} =$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \left[ \frac{-H_4(z)}{24(w_1 + w_2)} - \frac{(w_2 - w_1)^2 H_6(z)}{72(w_1 + w_2)^3} \right],$$

where $H_j(z)$ is the Hermite polynomial of order $j$. We obtain then

$$P_{w_1 w_2}(k) = \frac{1}{\sqrt{2\pi(w_1+w_2)}} \exp\left(-\frac{z^2}{2}\right) \Bigg\{ 1 + \frac{1}{6\sqrt{w_1+w_2}} \frac{w_2-w_1}{w_1+w_2}(z^3-3z) +$$

$$+ \frac{1}{24(w_1+w_2)} \left[ -z^4+6z^2+3+\frac{1}{3}\left(\frac{w_2-w_1}{w_1+w_2}\right)^2 (-z^6+15z^4-45z^2+15) \right] \Bigg\} +$$

$$+ o\left[\frac{1}{(w_1+w_2)^{\frac{3}{2}}}\right],$$

$$P_w(k) = \frac{1}{2\sqrt{\pi w}} \exp\left(-\frac{z^2}{2}\right) \left[ 1 + \frac{1}{48w}(-z^4+6z^2+3) \right] + o\left(\frac{1}{w^2}\right).$$

I express my gratitude to A. Rényi for his valuable suggestions.

(*Received 7 December 1954*)

## References

[1] J. O. Irwin, The frequency distribution of the difference between two independent variates following the same Poisson distribution, *Journal of the Royal Statistical Society*, (1937).

[2] M. Fisz, Rozkład graniczny zmiennej losowej będącej róznicą dwóch niezależnych zmiennych losowych o rozkładzie Poissona, *Zastosowania Matematyki*, 1 (1953), pp. 41 - 45.

[3] M. Fisz, The limiting distribution of some function of two independent random variables and its statistical application, *Colloquium Mathematicum*, 3 (1955), pp. 138—146.

[4] C. G. Esseen, Fourier analysis of distribution functions. A mathematical study of the Laplace—Gaussian law, *Acta Math.*, 77 (1945), pp. 1—125.

## УТОЧНЕНИЕ ОДНОЙ ВЕРОЯТНОСТНОЙ ПРЕДЕЛЬНОЙ ТЕОРЕМЫ И ЕЕ ПРИМЕНЕНИЕ К БЕССЕЛЕВЫМ ФУНКЦИЯМ

М. Ф и ш (Варшава)

(Р е з ю м е)

Применяя общую теорему Э с с е е н а (см [4]), автор дает асимптотическое разложение (7) для распределения вероятностей разности двух независимых пуассоновских случайных величин. С помощью этого разложения получается асимптотическая формула (13) для бесселевой функции порядка $k$ с чисто мнимым аргументом.

# BEMERKUNGEN ZUR APPROXIMATION EINER REELLEN ZAHL DURCH BRÜCHE

Von

P. SZÜSZ (Budapest)

(*Vorgelegt von* P. TURÁN)

Es sei $\alpha$ eine Irrationalzahl. Zu jeder $\alpha$ gibt es bekanntlich unendlich viele Zahlenpaare $p, q$, für die

(1) $$|q\alpha - p| < \frac{1}{q},$$

und sogar

(1') $$|q\alpha - p| < \frac{1}{\sqrt{5}q}$$

gilt; der Faktor $5^{-\frac{1}{2}}$ auf der rechten Seite von (1') läßt sich durch keinen kleineren ersetzen.

Während der Abfassung eines Lehrbuches hat Prof. TURÁN folgende Frage gestellt: Es sei $\delta$ eine reelle Zahl mit

$$0 \leqq \delta \leqq 1$$

und man betrachte die Zahlenpaare $p, q$ für die

(2) $$|q\alpha - p| < q^{\delta-1}$$

gilt. Was läßt sich dann über die „Dichtigkeit" der $q$ aussagen? Wenn also $M(\alpha, \delta, n)$ die Anzahl der $q \leqq n$ bedeutet, für die (2) gültig ist, so soll das Verhalten des Quotienten $\dfrac{M(\alpha, \delta, n)}{n}$ untersucht werden. Es soll ferner untersucht werden, was sich über die „Lokalisation" dieser $q$ aussprechen läßt; es soll also für jedes $n$ eine möglichst kleine Zahl $m = m(n)$ derart angegeben werden, daß zwischen $n$ und $n+m$ mindestens ein (2) erfüllendes $q$ existiert.

Prof. TURÁN hat bewiesen, daß stets

(3) $$M(\alpha, \delta, n) > n^{\delta} \qquad (\delta > 0)$$

gilt, ferner, daß ebenfalls für $\delta > 0$, zwischen $n$ und $n^{\frac{1}{\delta}}$ mindestens ein (2) erfüllendes $q$ liegt.

Prof. TURÁN hat seine obigen Resultate P. ERDÖS und V. JARNIK brieflich mitgeteilt. In seinem Antwortsbriefe hat P. ERDÖS folgendes bewiesen: Es gilt

(4) $$M(\alpha, 0, n) > K \log n \qquad (K \text{ konstant}),$$

sogar etwas schärfer: bezeichnet $M_c(\alpha, 0, n)$ die Anzahl der natürlichen $q \leqq n$ mit $|q\alpha - p| < cq^{-1}$, so gilt für $c \leqq 5^{-\frac{1}{2}}$ stets

(4′) $$M_c(\alpha, 0, n) > K_c \log n,$$

wobei zu beachten ist, daß $K_c$ nicht einmal von $\alpha$ abhängt.

V. JARNIK hat für die Sätze von TURÁN und auch für (4′) einen Beweis gefunden und die Frage aufgeworfen, ob sich (3) bzw. (4) nicht verbessern lassen.

In der vorliegenden Note wiedergebe ich mit der freundlichen Erlaubnis von Prof. TURÁN seinen Beweis für (3), zeige, daß (3) im allgemeinen nicht verbessert werden kann und gebe die möglichst beste Lokalisation der (2) erfüllenden $q$ an. Daß (4) scharf ist, kann z. B. bei $\alpha = \frac{\sqrt{5}-1}{2}$ durch eine leichte Rechnung bestätigt werden. Ich beweise die folgenden Sätze:

SATZ 1.[1] *Für $\delta > 0$ gilt* (3); *andererseits gibt es ein $\alpha$, für welches mit unendlich vielen $n$*

(5) $$M(\alpha, \delta, n) < K(\delta) n^{\delta}$$

*gilt, wobei $K(\delta)$ nur von $\delta$ abhängt.*

SATZ 2. *Es sei $\varepsilon > 0$ und $0 \leqq \delta \leqq 1$ gegeben. Dann gibt es ein nur von $\varepsilon$ und $\delta$ abhängiges $N_0 = N_0(\varepsilon, \delta)$ so, daß zu einem reellen $\alpha$ zwischen $N$ und $\varepsilon N^{1+\delta}$ stets ein $n$ mit*

$$|n\alpha - m| < n^{-\delta} \qquad (m \text{ eine geeignete ganze Zahl})$$

*liegt, falls nur $N > N_0$.*

SATZ 3. *Es sei $g(x)$ eine positive, für $x \to \infty$ monoton gegen Null strebende Funktion. Dann existiert eine Irrationalzahl $\alpha$ und unendlich viele $N_r (r = 1, 2, \ldots)$ für die zwischen*

$$N_r \quad \text{und} \quad g(N_r) N^{1+\delta}$$

*kein $n$ existiert, für welches mit irgendeinem ganzen $n$*

(2′) $$|n\alpha - m| < n^{-\delta}$$

*gilt.*

Dieser Satz besagt, daß Satz 2 nicht mehr verschärft werden kann.

Bevor ich zum Beweis der ausgesprochenen Tatsachen übergehe, möchte ich bemerken, daß nichts ausgesprochen werden kann, wenn man außer (2)

[1] Wie mir Prof. TURÁN mitteilte, hat diese Ungleichung P. ERDÖS auch gefunden.

bzw. $|n\alpha - m| < n^{-\delta}$ noch $(p, q) = 1$ bzw. $(n, m) = 1$ verlangt. Dies folgt leicht aus den elementaren Tatsachen der Kettenbruchlehre.[2]

Nun beweise ich die ausgesprochenen Sätze.

BEWEIS DES SATZES 1.

Den folgenden Beweis von (3) hat mir Prof. TURÁN mitgeteilt.

Man setze $k = \left[n^{\frac{1}{1+\delta}}\right]$, $r = \left[n^{\frac{\delta}{1+\delta}}\right]$ und man betrachte die Zahlen

$$m_\varkappa = \varkappa\alpha - [\varkappa\alpha] \qquad (\varkappa = 0, 1, \ldots, rk).$$

Nach dem Schubfachprinzip gibt es ein Intervall

$$I_l = \left(\frac{l}{k}, \frac{l+1}{k}\right)$$

derart, daß in $I_l$ mindestens $r+1$ Zahlen $m_{\varkappa_\varrho}$ $(\varrho = 1, 2, \ldots, r+1)$ liegen; es gilt also

$$|m_{\varkappa_\varrho} - m_{\varkappa_1}| = |y\alpha - x| < \frac{1}{k} \qquad (\varrho = 2, 3, \ldots, r+1), \tag{6}$$

wobei $y$ eine natürliche Zahl $\leqq kr$ bedeutet, während $x$ ganz ist. Aus (6) und aus der Definition von $k$ und $r$ folgt, daß für mindestens $\left[n^{\frac{\delta}{1+\delta}}\right]$ Zahlen unter $kr$, also erst recht unter $n$ die Relation

$$|y\alpha - x| < \frac{1}{k} \leqq \left[y^{\frac{1}{1+\delta}}\right]^{-1} \tag{7}$$

gilt. Bestimmt man $\delta'$ aus

$$\frac{1}{1+\delta} = 1 - \delta',$$

so erhält man (3) mit $\delta'$ statt $\delta$, was zu beweisen war.

Nun beweise ich (5). Es bezeichnen (hier und im Folgenden) $a_0, a_1, \ldots$ die Kettenbruchnenner von $\alpha$, $\frac{A_0}{B_0}, \frac{A_1}{B_1}, \ldots$ die Näherungsbrüche von $\alpha$. Ferner bezeichne $N(\alpha, I, n)$ die Anzahl der $\nu \leqq n$ mit $(\nu\alpha) \in I$, wobei $I$ ein gegebenes Teilintervall von $(0,1)$ bezeichnet, während $(\ )$ den Bruchteil der in der Klammer stehenden Zahl bedeutet. Im Folgenden bezeichne $I_n$ der Kürze halber die Vereinigung der Intervalle $(0, n^{\delta-1})$ und $(1 - n^{\delta-1}, 1)$. Dies wird, da wir $\delta$ $(\delta > 0)$ festgelegt haben, kein Mißverständnis verursachen. Dann gilt offenbar

$$M(\alpha, \delta, n) < \sum_{\nu=1}^{r+1} \{N(\alpha, I_{B_{\nu-1}}, B_\nu) - N(\alpha, I_{B_{\nu-1}}, B_{\nu-1})\}, \tag{8}$$

[2] Prof. TURÁN hat die entsprechenden Fragen auch im Falle simultaner Approximation mehrerer reellen Zahlen aufgeworfen. Die entsprechenden Resultate haben auch Anwendungen. Auf diese Fragen beabsichtige ich in einer späteren Note zurückzukommen.

wobei $r$ denjenigen Index bedeutet, für welchen $B_r \leqq n < B_{r+1}$ gilt. Die Rolle der $B_\nu$ kann in (8) jede streng zunehmende Folge ganzer Zahlen übernehmen. Daß wir eben die Folge der Näherungsnenner von $\alpha$ gewählt haben, wird später ausgenutzt werden, vgl. Formel (9).

Aus

$$|l\alpha - k| < l^{\delta-1} \qquad (l \text{ natürlich, } k \text{ ganz})$$

und

$$B_{\nu-1} \leqq l < B_\nu$$

folgt nämlich jedenfalls

$$(l\alpha) \in I_{B_{\nu-1}}$$

Man setze insbesondere $a_0 = a_1 = \cdots = 1$ $\left(\text{also } \alpha = \frac{1+\sqrt{5}}{2}\right)$, und $n = B_r$. Dann folgt aus (8)

$$M(\alpha, \delta, B_r) < \sum_{\nu=1}^{r+1} \{N(\alpha, I_{B_{\nu-1}}, B_\nu) - N(\alpha, I_{B_{\nu-1}}, B_{\nu-1})\}. \tag{8'}$$

Nach einem Satz von BEHNKE[3] gilt

$$\begin{cases} N(\alpha, I_{B_{\nu-1}}, B_\nu) = 2B_\nu B_{\nu-1}^{\delta-1} + O(1), \\ N(\alpha, I_{B_{\nu-1}}, B_{\nu-1}) = 2B_{\nu-1}^{\delta} + O(1). \end{cases} \tag{9}$$

Daher ist die rechte Seite von (8′) gleich

$$2\sum_{\nu=1}^{r+1} B_{\nu-1}^{\delta-1}(B_\nu - B_{\nu-1}) + O(r) = 2\sum_{\nu=1}^{r+1} B_{\nu-1}^{\delta}\left(\frac{B_\nu}{B_{\nu-1}} - 1\right) + O(r),$$

also $\left(\text{wegen } \alpha = \frac{\sqrt{5}+1}{2},\ \frac{B_\nu}{B_{\nu-1}} < 2\right)$ kleiner als

$$2\sum_{\nu=1}^{r+1} B_{\nu-1}^{\delta} + O(\log B_{r+1}).$$

Da (für jedes $\alpha$) bekanntlich

$$B_{r+1} > 2B_{r-1} \qquad (r = 1, 2, \ldots)$$

gilt, folgt hieraus

$$M(\alpha, \delta, B_r) < 4B_r^{\delta}(1 + 2^{-\delta} + 2^{-2\delta} + \cdots) + O(\log B_{r+1}) =$$
$$= \frac{4B_r^{\delta}}{1 - 2^{-\delta}} + O(\log B_r),$$

womit (5) bewiesen ist. Die Frage nach dem kleinsten Wert $K_0(\delta)$ von $K(\delta)$, d. h. nach derjenigen Zahl, für die

$$M(\alpha, \delta, n) > K_0(\delta) n^{\delta},$$

[3] H. BEHNKE, *Hamburger Abhandlungen*, 3 (1924), S. 261—318.

aber bei beliebigem $\varepsilon > 0$ für geeignetes $\alpha$ und unendlich viele $n$

$$M(\alpha, \delta, n) < (K_0(\delta) + \varepsilon) n^\delta$$

gilt, bleibt offen. Sowohl der Turánsche Beweis von (3), als mein Ansatz, der zu (5) führte, scheint in dieser Hinsicht zu roh zu sein.

Beweis des Satzes 2. $a_0, a_1, \ldots$ bzw. $\frac{A_0}{B_0}, \frac{A_1}{B_1}, \ldots$ habe dieselbe Bedeutung, wie bei dem Beweis von (5). Bekanntlich gilt für jedes natürliche $r$

$$(10) \qquad |B_r \alpha - A_r| < \frac{1}{B_{r+1}},$$

also für jedes natürliche $c$

$$(10') \qquad |cB_r \alpha - cA_r| < \frac{c}{B_{r+1}}.$$

Für $c \leqq \left(\frac{B_{r+1}}{B_r^\delta}\right)^{\frac{1}{1+\delta}}$ gilt $\frac{c}{B_{r+1}} \leqq \frac{1}{(cB_r)^\delta}$, wegen (10') ist also für natürliches

$$c \leqq \left(\frac{B_{r+1}}{B_r^\delta}\right)^{\frac{1}{1+\delta}}$$

erst recht

$$(11) \qquad |cB_r \alpha - cA_r| < (cB_r)^{-\delta}.$$

Es sei nun $n$ gegeben, und es bezeichne $r = r(n)$ denjenigen Index, für welchen

$$(12) \qquad B_r \leqq n < B_{r+1}$$

gilt.

Es darf ohne Beschränkung der Allgemeinheit

$$(13) \qquad B_{r+1} > 4B_r,$$

also $a_{r+1} \geqq 4$ angenommen werden; sonst wäre nämlich $B_{r+1}$ eine Zahl, die (2) genügt und für die $n < B_{r+1} \leqq 4n$.

Wir unterscheiden nun zwei Fälle.

1) Es sei zuerst

$$(14) \qquad n \leqq (B_r B_{r+1})^{\frac{1}{1+\delta}} - B_r.$$

Für jedes natürliche $c \leqq \left(\frac{B_{r+1}}{B_r^\delta}\right)^{\frac{1}{1+\delta}}$ gilt nach (11)

$$(15) \qquad |cB_r \alpha - cA_r| < (cB_r)^{-\delta}.$$

Je zwei konsekutive der Zahlen $cB_r\left(c=1,2,\ldots,\left[\left(\frac{B_{r+1}}{B_r^\delta}\right)^{\frac{1}{1+\delta}}\right]\right)$ haben voneinander den Abstand $B_r$; wegen (14) ist jedenfalls

$$n \leqq B_r\left[\left(\frac{B_{r+1}}{B_r^\delta}\right)^{\frac{1}{1+\delta}}\right].$$

Daher gibt es ein $cB_r$ mit (15), für welches $0 \leqq cB_r - n < B_r$, womit unser Satz 2 im Falle 1) bewiesen ist.

2) Es sei nun

$$n > (B_r B_{r+1})^{\frac{1}{1+\delta}} - B_r, \tag{16}$$

also wegen (13) und $\delta \leqq 1$ gewiß

$$n > \left(\frac{B_{r+1}B_r}{4}\right)^{\frac{1}{1+\delta}}. \tag{16'}$$

Dann folgt aus (12) und (16')

$$n < B_{r+1} < \frac{4}{B_r} n^{1+\delta};$$

da $B_{r+1}$, als Näherungsnenner von $\alpha$ der Relation $|B_{r+1}\alpha - A_{r+1}| < B_{r+1}^{-1}$ genügt und $B_r \to \infty$ für $r \to \infty$, ist unser Satz 2 auch im Falle 2) bewiesen.

Beweis des Satzes 3. Es bezeichnen $a_0, a_1, \ldots$ die Kettenbruchnenner, $\frac{A_0}{B_0}, \frac{A_1}{B_1}, \ldots$ die Näherungsbrüche von $\alpha$.

Bekanntlich[4] gilt

$$|y\alpha - x| \geqq |B_{r-1}\alpha - A_{r-1}|, \tag{17}$$

falls $y$ natürlich, $x$ ganz ist und $y < B_r$ besteht.

Es sei $n$ gegeben; $r$ bedeutet denjenigen Index, für den (12) gilt. Dann läßt $n$ eine Darstellung von der Form

$$n = qB_r + s \qquad (1 \leqq q \leqq a_{r+1},\ 0 \leqq s < B_r) \tag{18}$$

zu. Ist dann $m$ eine beliebige ganze Zahl, so gilt

$$|n\alpha - m| = |qB_r\alpha + s\alpha - m| = |q(B_r\alpha - A_r) + s\alpha + qA_r - m|.$$

Ist insbesondere $s \geqq 1$ und $q \leqq \frac{a_{r+1}}{2}$, so folgt aus der bekannten Formel[5]

$$\left|\frac{B_r\alpha - A_r}{B_{r-1}\alpha - A_{r-1}}\right| = \frac{1}{\zeta_{r+1}} < \frac{1}{a_{r+1}}$$

[4] Vgl. z. B. Perron, *Die Lehre von den Kettenbrüchen*, II. Aufl. (Berlin—Leipzig, 1929), S. 52.

[5] Vgl. Perron, loc. cit., S. 42, Formel (4).

(hier bedeuten $\zeta_0, \zeta_1, \ldots$ die vollständigen Quotienten der Kettenbruchentwicklung von $\alpha$)

$$(19)\left\{\begin{aligned}&|n\alpha-m| \geqq |s\alpha-m+qA_r|-q|B_r\alpha-A_r| \geqq |B_{r-1}\alpha-A_{r-1}|-q|B_r\alpha-A_r| > \\ &\quad > |B_{r-1}\alpha-A_{r-1}|-\frac{q}{a_{r+1}}|B_{r-1}\alpha-A_{r-1}| = \\ &\quad = \left(1-\frac{q}{a_{r+1}}\right)|B_{r-1}\alpha-A_{r-1}| \geqq \frac{1}{2}|B_{r-1}\alpha-A_{r-1}|.\end{aligned}\right.$$

Alles bisher bewiesene gilt allgemein. Wir versuchen nun, die Kettenbruchnenner $a_0, a_1, \ldots$ von $\alpha$ geeignet zu wählen.

Ich nehme an, daß die $a_r$ $(r=0, 1, \ldots)$ so stark anwachsen, daß stets

$$(20)\qquad B_{r+1} > 4^{\frac{1+\delta}{\delta}} B_r^{\frac{1}{\delta}}$$

gilt. Dann behaupte ich, daß für $r=1, 2, \ldots$ zwischen

$$2(B_r B_{r+1})^{\frac{1}{1+\delta}} \quad \text{und} \quad \frac{B_{r+1}}{2}$$

*keine* solche natürliche Zahl $n$ liegt, für welche

$$|n\alpha-m| < n^{-\delta} \qquad (m \text{ ganz})$$

gilt.

Man betrachte nämlich für $2(B_r B_{r+2})^{\frac{1}{1+\delta}} < n < \frac{B_{r+1}}{2}$ erfüllendes $n$ und ganzes $m$ den Ausdruck

$$(21)\qquad |n\alpha-m|.$$

Wir stellen $n$ in der Form (18) dar; es ist also eine untere Schranke für den Ausdruck

$$(21')\qquad |qB_r\alpha+s\alpha-m|$$

anzugeben, wobei $m$ sämtliche ganze Zahlen durchläuft, während

$$(22)\qquad 2(B_r B_{r+1})^{\frac{1}{1+\delta}} < qB_r+s < \frac{B_{r+1}}{2}$$

besteht. Wir unterscheiden nun zwei Fälle, je nach dem $s=0$ oder $s \geqq 1$ gilt.

1) Es sei zuerst $s=0$.

Dann ist wegen (22)

$$(23)\qquad 2\left(\frac{B_{r+1}}{B_r^{\delta}}\right)^{\frac{1}{1+\delta}} < q < \frac{B_{r+1}}{2B_r} < \frac{a_{r+1}}{2}+\frac{1}{2}.$$

Ich zeige, daß für $q \leqq \frac{a_{r+1}}{2}$

$$|qB_r\alpha - m| > \frac{2}{n^\delta}$$

gilt. Dies kann folgendermaßen geschehen: Ist $|qB_r\alpha - m| < 1$, so ist jedenfalls entweder

$$m = [qB_r\alpha]$$

oder

$$m = [qB_r\alpha] + 1.$$

Daher gilt

$$|qB_r\alpha - m| > \min\{(qB_r\alpha), 1 - (qB_r\alpha)\}.$$

Nun ist aber für $q \leqq \frac{a_{r+1}}{2}$

$$\min\{(qB_r\alpha), 1 - (qB_r\alpha)\} = q|B_r\alpha - A_r|.$$

Daher ist wegen der bekannten Relation $|B_r\alpha - A_r| > \frac{1}{2}\frac{1}{B_{r+1}}$ und wegen (23)

$$|qB_r\alpha - m| \geqq q|B_r\alpha - A_r| > \frac{1}{2}\frac{q}{B_{r+1}} \geqq \left(\frac{B_{r+1}}{B_r^\delta}\right)^{\frac{1}{1+\delta}} \frac{1}{B_{r+1}} = (B_r B_{r+1})^{-\frac{\delta}{1+\delta}} > \frac{2}{n^\delta},$$

was zu beweisen war.

2) Es sei jetzt in (21′)

$$1 \leqq s < B_r.$$

Die zweite Ungleichung der Formel (23) gilt auch in diesem Falle. Daher folgt aus (19) und (20)

$$|n\alpha - m| > \frac{1}{2}|B_{r-1}\alpha - A_{r-1}| > \frac{1}{4B_r} > (B_r B_{r+1})^{-\frac{\delta}{1+\delta}} > \frac{2}{n^\delta}, \tag{24}$$

womit wir unsere Behauptung auch im Falle 2) bewiesen haben.

Wir haben also gezeigt, daß falls die Kettenbruchnenner $a_0, a_1, \ldots$ von $\alpha$ genügend stark anwachsen, zwischen

$$2(B_r B_{r+1})^{\frac{1}{1+\delta}}$$

und

$$\frac{1}{2}B_{r+1}$$

*keine* Zahl $n$ liegt, die mit irgendeinem ganzen $m$

$$|n\alpha - m| < n^{-\delta} \tag{25}$$

leistet.

Setzt man $2(B_r B_{r+1})^{\frac{1}{1+\delta}} = N_r$, so besagt unser Satz, daß zwischen

(26) $$N_r \quad \text{und} \quad \frac{N_r^{1+\delta}}{2^{2+\delta} B_r}$$

kein $n$ mit (25) liegt.

Schreibt man

$$f(x) = \frac{1}{B_r} \quad \text{für} \quad B_r \leqq x < B_{r+1}$$

und ist $g(x)$ eine positive, monoton gegen Null strebende, aber sonst beliebige gegebene Funktion, so kann durch genügend starkes Anwachsen der $a_r$ offenbar erreicht werden, daß für unendlich viele $r$

$$f(N_r) > g(N_r)$$

gilt. Hieraus und aus (26) folgt, daß für unendlich viele $r$ zwischen

$$N_r \quad \text{und} \quad g(N_r) N_r^{1+\delta}$$

kein $n$ mit (25) liegt. Damit haben wir auch Satz 3 bewiesen.

Zum Schluß spreche ich Herrn Professor TURÁN meinen Dank aus dafür, daß er mich auf diese Probleme aufmerksam gemacht hat.

BUDAPEST, INSTITUT FÜR ANGEWANDTE MATHEMATIK
DER UNGARISCHEN AKADEMIE DER WISSENSCHAFTEN

*(Eingegangen am 29. Januar 1955.)*

## ЗАМЕЧАНИЯ К АППРОКСИМАЦИИ ВЕЩЕСТВЕННЫХ ЧИСЕЛ ДРОБЯМИ

П. Сюс (Будапешт)

(Резюме)

Пусть $\alpha$ — вещественное число, а $\delta$ вещественное число, удовлетворяющее условию $0 \leqq \delta < 1$. Рассмотрим натуральные числа $q$, для которых при некотором целом $p$ имеет место

(1) $$|q\alpha - p| < q^{-\delta}.$$

Обозначим через $M(\alpha, \delta, n)$ число тех значений $q$, не превышающих $n$, для которых выполняется (1). Тогда справедлива следующая

Теорема 1. Для всякого вещественного $\alpha$ имеет место

(2) $$M(\alpha, \delta, n) > n^{1-\delta},$$

и, с другой стороны, существует значение $\alpha$, для которого

(3) $$M(\alpha, \delta, n) < K(\delta) n^{1-\delta},$$

где число $K(\delta)$ зависит только от $\delta$. Его наименьшее значение неизвестно.

О „локализации“ натуральных чисел $q$, удовлетворяющих (1), можно сказать следующее:

Теорема 2. Для данного $\varepsilon > 0$ существует такое число $N = N(\varepsilon)$, что между

$$n \quad \text{и} \quad \varepsilon n^{1+\delta}$$

попадает хотя бы одно значение $q$, удовлетворяющее (1), если только $n > N$.

Теорема 3. Пусть $f(x)$ — положительная функция, стремящаяся к нулю для $x \to \infty$. Тогда существует такое число $\alpha$, что для бесконечного множества значений $n$ между

$$n \quad \text{и} \quad f(n) n^{1+\delta}$$

нет ни одного значения $q$, удовлетворяющего (1), так что теорема 2 не допускает улучшения.

Приведенное доказательство от (2) принадлежит профессору П. Туран. Все другие утверждения доказываются с помощью теории цепных дробей.

# EXTENSION OF FUNCTIONS SATISFYING A LIPSCHITZ CONDITION

By

J. CZIPSZER (Budapest) and L. GEHÉR (Miskolc)

(*Presented by* G. ALEXITS)

It is a simple fact that a function $f(x)$ defined on a set of the real line and satisfying a Lipschitz condition

$$|f(x)-f(y)| \leqq K|x-y|$$

may be extended to the whole real line so that the Lipschitz condition with the same constant $K$ remains valid for the extended function. The present paper deals with the generalizations of this theorem for functions defined on a subset of a given metric space.

## § 1

In this section we shall prove the following theorem:

THEOREM I. *Let $X$ be a metric space. Let $f(x)$ be a real function defined on a subset $S$ of $X$ and satisfying a Lipschitz condition of the form*

$$|f(x)-f(y)| \leqq K\overline{xy}$$

*($K \geqq 0$, $\overline{xy}$ denotes the distance of $x$ and $y$), i. e.*

$$f(x) \in \mathrm{Lip}_S(1, K).$$

*Then there is a real function $F(x)$ defined on $X$, equal to $f(x)$ on $S$ and satisfying the same Lipschitz condition, i. e.*

$$F(x) \in \mathrm{Lip}_X(1, K).$$

If $X$ is the $n$-dimensional Euclidean space $R_n$, then this theorem is contained in a theorem of M. D. KIRSZBRAUN,[1] which asserts that if $f(x)$ is

[1] M. D. KIRSZBRAUN, Über die zusammenziehende und Lipschitzsche Transformation, *Fundamenta Math.*, **22** (1934), pp. 77—108 (esp. p. 94, Hauptsatz I).

defined in a subspace of $R_n$, takes values in $R_n$ and satisfies the Lipschitz condition

$$\overline{f(x)f(y)} \leqq K\overline{xy},$$

then it may be extended to the whole space $R_n$ preserving the same Lipschitz condition.

From the other hand, Theorem I guarantees the possibility to extend every function, defined on a subset of a metric space, ranging over $R_n$ (or more generally over any $n$-dimensional Banach space) and satisfying a Lipschitz condition with constant $K$, to a function which is defined throughout the space and satisfies a Lipschitz condition with constant $\sqrt{n}K$ (or with some other constant).

We shall show that the function

$$F(x) = \inf_{y\in S} [f(y) + K\overline{xy}]$$

is the desired extension of $f(x)$.

Before starting the proof, we should like to show how one can hit upon this formula. If $x$ is any point outside $S$, then $f(x)$ must be defined so that the inequalities

$$f(y) - K\overline{xy} \leqq f(x) \leqq f(y) + K\overline{xy} \tag{1}$$

shall hold for every $y \in S$. Now for any two points $y, z$ of $S$

$$f(y) - f(z) \leqq K\overline{yz} \leqq K\overline{xy} + K\overline{xz}$$

and by this

$$\sup_{y\in S} [f(y) - K\overline{xy}] \leqq \inf_{z\in S} [f(z) + K\overline{xz}].$$

This inequality shows that with suitable choice of $f(x)$ we can satisfy (1): with

$$\sup_{y\in S} [f(y) - K\overline{xy}] \leqq f(x) \leqq \inf_{y\in S} [f(y) + K\overline{xy}] \tag{2}$$

this is already attained. Now repeating this procedure we can extend $f(x)$ by every step to one new point, and with the help of the well-ordering theorem we can define $f(x)$ at last everywhere. This procedure is nearly the same as the one used for the proof of the theorem of Hahn—Banach.[2] But fortunately this successive construction can be replaced by a much more simple construction: if we define $f(x)$ for every $x$ to be equal to the left or right end-point of the interval (2), we get a function which satisfies the prescribed Lipschitz condition not only "between" two points one being in $S$ and the other outside $S$, but "between" any two points of the space.

Proof of theorem I. A) $F(x)$ is everywhere defined in $X$. For a fixed $z \in X$ we have the inequality

$$f(y) + K\overline{xy} = f(z) + K\overline{xy} + [f(y) - f(z)] \geqq f(z) + K\overline{xy} - K\overline{yz} \geqq f(z) - K\overline{xz}$$

which shows that the set of values

$$f(y) + K\overline{xy} \quad y \in S$$

for any fixed $x$ is bounded from below and that gives the assertion.

[2] S. Banach, *Théorie des opérations linéaires* (Warszawa, 1932), Théorème I, Chap II and Théorème 2, Chap. IV.

B) If $x \in S$ and $y \in S$, then

$$f(x)+K\overline{xx}=f(x)$$

and

$$f(y)+K\overline{xy} \geqq f(x).$$

Consequently,

$$F(x)=f(x) \quad \text{for} \quad x \in S$$

and so $F(x)$ is in fact an extension of $f(x)$.

C) We shall now prove that $F(x) \in \mathrm{Lip}_X(1, K)$. Let $x_1$ and $x_2$ be two points of $X$. For a given $\varepsilon > 0$ we may choose a $y \in S$ so that

$$F(x_1) \geqq f(y)+K\overline{x_1 y}-\varepsilon$$

holds. Then we have

$$F(x_2)-F(x_1) \leqq f(y)+K\overline{x_2 y}-[f(y)+K\overline{x_1 y}-\varepsilon]=$$
$$=K(\overline{x_2 y}-\overline{x_1 y})+\varepsilon \leqq K\overline{x_1 x_2}+\varepsilon.$$

We get similarly

$$F(x_1)-F(x_2) \leqq K\overline{x_1 x_2}+\varepsilon.$$

Hence

$$|F(x_1)-F(x_2)| \leqq K\overline{x_1 x_2}+\varepsilon.$$

Since $\varepsilon$ is arbitrary, we see that indeed

$$F(x) \in \mathrm{Lip}_X(1, K).$$

Thus the proof of Theorem I is complete.

## § 2

We shall now give a sort of generalization of Theorem I.

THEOREM II. *Let $X$ be a metric space. Let $f(x)$ be a real function defined on a closed subset $S$ of $X$ which satisfies in every point $x$ of $S$ a Lipschitz condition of the form*

$$|f(x)-f(y)| \leqq K\overline{xy} \quad \text{for} \quad \overline{xy} < \delta \qquad (K \geqq 0, \delta > 0),$$

*where $K$ and $\delta$ depend on $x$. (We shall denote this property by saying that $f(x)$ belongs to $\mathrm{LIP}_S$ or briefly to $\mathrm{LIP}$). Then there is a real function $F(x)$ defined on $X$, equal to $f(x)$ on $S$ and which belongs to $\mathrm{LIP}_X$.*

PROOF. For the time being let us suppose that $f(x)$ is bounded

$$|f(x)| \leqq M \quad \text{for} \quad x \in S.$$

It is obvious that by a suitable change of $K$ the Lipschitz condition

$$|f(x)-f(y)| \leqq K(x)\overline{xy}$$

shall hold for every $y$. The notation $K(x)$ sets in evidence that $K$ depends

on $x$. We shall show that the function

$$F(x)=\inf_{y\in S}[f(y)+K(y)\overline{xy}]$$

is the desired extension of $f(x)$.

It is obvious that — since $f(x)$ is bounded — $F(x)$ is defined everywhere. In the same way as in the previous §, we see that

$$F(x)=f(x) \quad \text{if} \quad x\in S.$$

We have still to prove that $F(x)$ satisfies a Lipschitz condition in every point $x$.

We choose a point $x_0$. First we suppose that $x_0$ belongs to $S$. Then we have

$$F(x)\leqq f(x_0)+K(x_0)\overline{xx_0}$$

or

$$F(x)-f(x_0)=F(x)-F(x_0)\leqq K(x_0)\overline{xx_0}.$$

If $y$ belongs to $S$ and $y\neq x_0$, then

$$f(y)-f(x_0)\geqq -K(y)\overline{x_0y}$$

and

$$f(y)-f(x_0)\geqq -K(x_0)\overline{x_0y}.$$

Multiplying the first inequality with $\dfrac{\overline{xy}}{\overline{xy}+\overline{xx_0}}$, the second with $\dfrac{\overline{xx_0}}{\overline{xy}+\overline{xx_0}}$ we get

$$f(y)-f(x_0)\geqq -[K(y)\overline{xy}+K(x_0)\overline{xx_0}]\frac{\overline{x_0y}}{\overline{xy}+\overline{xx_0}}\geqq -[K(y)\overline{xy}+K(x_0)\overline{xx_0}].$$

Hence

$$f(y)+K(y)\overline{xy}\geqq f(x_0)-K(x_0)\overline{xx_0},$$

and this last inequality is true also in the case $y=x_0$.

We see at once that

$$F(x)=\inf_{y\in S}[f(y)+K(y)\overline{xy}]\geqq f(x_0)-K(x_0)\overline{xx_0}$$

or

$$F(x)-f(x_0)\geqq -K(x_0)\overline{xx_0}.$$

So the inequality

$$|F(x)-f(x_0)|=|F(x)-F(x_0)|\leqq K(x_0)\overline{xx_0}$$

holds, i. e. $F(x)$ satisfies a Lipschitz condition in every point of $S$.

Now let us suppose that $x_0\in X-S$. We denote by $d=d(x_0,S)$ the (positive) distance of $x_0$ from $S$ and consider the open neighbourhood $U$ of $x_0$ with center $x_0$ and radius $\frac{1}{2}d$.

For $x\in U$ and for any fixed $z\in S$

$$F(x)\leqq f(z)+K(z)\overline{xz}\leqq f(z)+K(z)\,d(z,U).$$

From the other side

$$F(x) \geqq \inf_{y \in S} f(y) \geqq -M.$$

The last inequalities show that $|F(x)|$ is bounded in $U$ by some $N$:

$$|F(x)| \leqq N \qquad (x \in U).$$

Now let us choose a point $x_1$ from $U$. For a given $\varepsilon > 0$ we can find a $y \in S$ so that the inequality

(1) $$F(x_1) \geqq f(y) + K(y)\overline{x_1 y} - \varepsilon$$

shall hold. If we subtract this inequality from

$$F(x_0) \leqq f(y) + K(y)\overline{x_0 y},$$

we get

(2) $$F(x_0) - F(x_1) \leqq K(y)(\overline{x_0 y} - \overline{x_1 y}) + \varepsilon \leqq K(y)\overline{x_0 x_1} + \varepsilon.$$

From (1) it follows

$$K(y) \leqq \frac{F(x_1) - f(y) + \varepsilon}{\overline{x_1 y}} \leqq \frac{N + M + \varepsilon}{\frac{d}{2}}.$$

Hence and by (2)

$$F(x_0) - F(x_1) \leqq \frac{N + M + \varepsilon}{\frac{d}{2}} \overline{x_0 x_1} + \varepsilon,$$

and since $\varepsilon$ is arbitrary, we have

$$F(x_0) - F(x_1) \leqq 2\frac{N+M}{d} \overline{x_0 x_1}.$$

The analogous inequality

$$F(x_1) - F(x_0) \leqq 2\frac{N+M}{d} \overline{x_0 x_1}$$

may be deduced in the same way, so that together we have

$$|F(x_1) - F(x_0)| \leqq 2\frac{N+M}{d(x_0, S)} \overline{x_0 x_1}$$

for

$$\overline{x_0 x_1} < \frac{1}{2} d(x_0, S),$$

i. e. $F(x)$ satisfies a Lipschitz condition in every point of $X - S$.

Thus Theorem II is proved under the hypothesis of the boundedness of $f(x)$.

In the case the function which is to be extended is not necessarily bounded we shall need the following remark:

If $|f(x)|<M$ on $S$, then the extension of $f(x)$ may be chosen so that this inequality will hold throughout $X$. We can assume that the Lipschitz constants $K(x)$ are all larger then 1. Then for $x \in X-S$

$$F(x) \geqq -M+d(x, S) > -M.$$

The function

$$F_1(x) = \min [F(x), M]$$

is obviously also an extension of $f(x)$ which belongs to $\mathrm{LIP}_X$; moreover $F_1(x)$ satisfies the inequality

$$-M < F_1(x) \leqq M.$$

In the same way we can get a function $F_2(x)$ defined on $X$ which is an extension of $-f(x)$, belongs to $\mathrm{LIP}_X$ and satisfies the inequality

$$-M < F_2(x) \leqq M.$$

Now the function $\Phi(x) = \frac{1}{2}[F_1(x) - F_2(x)]$ belongs also to $\mathrm{LIP}_X$, is also an extension of $f(x)$ and satisfies the inequality

$$-M < \Phi(x) < M.$$

Hence our statement is proved.

Now let us depart of the hypothesis of the boundedness of $f(x)$. Instead of $f(x)$ let us consider the function

$$g(x) = \operatorname{arctg} f(x)$$

which obviously also belongs to $\mathrm{LIP}_S$.

By what we have already proved, we know that there exists a function $G(x)$ identical with $g(x)$ on $S$ and which belongs to $\mathrm{LIP}_X$. Moreover, we know that since $|g(x)| < \frac{\pi}{2}$, $G(x)$ may be chosen so that the inequality

$$|G(x)| < \frac{\pi}{2}$$

shall hold everywhere. Now if we put

$$F(x) = \operatorname{tg} G(x),$$

then we have the desired extension of $f(x)$ since $\operatorname{tg} t$ is differentiable in $-\frac{\pi}{2} < t < \frac{\pi}{2}$.

## § 3

We should like to point out that Theorems I and II remain valid if the Lipschitz condition is replaced by the more general condition

$$|f(x)-f(y)| \leqq K\omega(\overline{xy}) \quad \text{resp.} \quad |f(x)-f(y)| \leqq K(x)\omega(\overline{xy}) \quad \text{for} \quad \overline{xy} < \delta(x)$$
$$(\delta(x) > 0),$$

where $\omega(t)$ is a function defined in $0 \leqq t < \infty$ with the following properties:

$\omega(t)$ is continuous and monotone increasing,
$\omega(0) = 0$, $\omega(t) > 0$ if $t > 0$,
$\omega(x+y) \leqq \omega(x) + \omega(y)$.

(An important example of such an $\omega(t)$ is $t^{\alpha}(0 < \alpha \leqq 1)$.) To see this, we have only to redefine the metric of $X$ by putting

$$\varrho(x, y) = \omega(\overline{xy}) \qquad (x, y \in X),$$

where $\varrho(x, y)$ designes the distance of $x$ and $y$ according to the new metric. The prescribed properties of $\omega(t)$ guarantee that $\varrho(x, y)$ satisfy all the axioms valid for a metric in a metric space. In this new metric $f(x)$ satisfies a Lipschitz condition

$$|f(x)-f(y)| \leqq K\varrho(x, y) \text{ resp. } |f(x)-f(y)| \leqq K(x)\varrho(x, y) \text{ for } \varrho(x, y) < \omega[\delta(x)].$$

Thus we have reduced this general case to the special one discussed above. We have still to remark that in case of Theorem II the set $S$ remains closed in the new metric.

*(Received 8 February 1955)*

**Note added in proof.** (10 June 1955.) Theorem I and its generalization to a Lipschitz condition of arbitrary exponent $p$ $(0 < p \leqq 1)$ with almost the same proof as in the present paper is to be found in a book of STEFAN BANACH, *Wstep do teorii funkcji rzeczywistych* (Warszawa—Wrocław, 1951), pp. 121—122. We were informed of this fact by an oral communication of Prof. STANISŁAW MAZUR late after the proof had been read.

## О РАСПРОСТРАНЕНИИ ФУНКЦИЙ, УДОВЛЕТВОРЯЮЩИХ УСЛОВИЯМ ЛИПШИЦА

Я. Ципсер (Будапешт) и Л. Гехер (Мишкольц)

(Резюме)

Пусть $\omega(t)$ $(0 \leqq t < \infty)$ — непрерывная монотонно возрастающая функция, удовлетворяющая условиям

$$\omega(0)=0, \quad \omega(t)>0 \quad (t>0), \quad \omega(x+y) \leqq \omega(x)+\omega(y).$$

(Например, $\omega(t)=t^{\alpha}$ $(0<\alpha \leqq 1)$.)

В настоящей работе доказываются следующие две теоремы:

Теорема 1. Пусть $f(x)$ — вещественная функция, определенная на подмножестве $S$ метрического пространства $X$ и удовлетворяющая условию

$$|f(x)-f(y)| \leqq \omega(\overline{xy}),$$

где $x$ и $y$ две произвольные точки $S$, а $\overline{xy}$ означает расстояние этих двух точек. Тогда существует функция $F(x)$, определенная на всем пространстве $X$ и совпадающая на $S$ с $f\ x$, для которой

$$|F(x)-F(y)| \leqq \omega(\overline{xy}),$$

где $x$ и $y$ две произвольные точки $X$.

Теорема 2. Пусть $f(x)$ — вещественная функция, определенная на замкнутом подмножестве $S$ метрического пространства $X$, и удовлетворяющая во всякой точке $x$ множества $S$ условию

$$|f(y)-f(x)| \leqq k(x)\,\omega(\overline{xy}) \qquad (\overline{xy}<\delta(x)),$$

где $y$ произвольная точка $S$, а $k(x)$ и $\delta(x)$ положительные числа, зависящие от $x$. Тогда можно найти функцию $F(x)$ определенную на $X$ и совпадающую на $S$ с $f(x)$, которая во всякой точке $x$ пространства $X$ удовлетворяет условию

$$|F(y)-F(x)| \leqq K(x)\,\omega(\overline{xy}) \qquad (\overline{xy}<\Delta(x)),$$

где $y$ произвольная точка $X$, а $K(x)$ и $\Delta(x)$ положительные числа, зависящие от $x$.

# ÜBER DAS GLIEDWEISE DIFFERENZIEREN EINER ORTHOGONALEN POLYNOMREIHE

Von

G. FREUD (Budapest)

(*Vorgelegt von* G. ALEXITS)

Es sei $p_n(x)$ die Folge der normierten Orthogonalpolynome mit der nichtnegativen Gewichtsfunktion $w(x) \in L$ in $(-1, +1)$. Wir bilden die Orthogonalentwicklung

$$f(x) \sim \sum_{\nu=0}^{\infty} c_\nu p_\nu(x) \tag{1}$$

einer Funktion $f(x)$, welche in $[-1, +1]$ $k$-mal stetig differenzierbar ist. Wir wollen in vorliegender Arbeit die absolute Konvergenz der Reihe

$$f^{(k)}(x) \sim \sum_{\nu=0}^{\infty} c_\nu p_\nu^{(k)}(x) \tag{2}$$

untersuchen, die durch formales gliedweise Differenzieren aus (1) entsteht.

Zuerst einige Bemerkungen. Betrachten wir die Fourierreihe

$$F(x) \sim \sum_{\nu=0}^{\infty} (a_\nu \cos \nu x + b_\nu \sin \nu x)$$

einer $k$-mal stetig differenzierbaren Funktion $F(x)$, dann erhält man durch $k$-maliges Differenzieren die Fourierreihe der Derivierten $F^{(k)}(x)$:

$$F^{(k)}(x) \sim \sum_{\nu=0}^{\infty} \nu^k \left[ a_\nu \cos \nu \left( x + \frac{k\pi}{2} \right) + b_\nu \sin \nu \left( x + \frac{k\pi}{2} \right) \right].$$

Die Frage der absoluten Konvergenz der derivierten Reihe ist also identisch mit der Frage der absoluten Konvergenz der Fourierreihe der Derivierten, die schon ausführlich behandelt wurde. Jedoch ist die Polynomreihe (2) (mit ev. Ausnahme bei speizieller Wahl von $w(x)$) keine Orthogonalreihe, also kann man ohne weiteres die bei der Untersuchung von Orthogonalreihen üblichen Methoden (vgl. z. B. STEČKIN [5]), ja sogar nicht einmal die Resultate des Verfassers (G. FREUD [3]) bezüglich der absoluten Konvergenz von Orthogonalpolynomreihen, auf die Reihe (2) nicht übertragen. Allerdings stellt es sich heraus, daß (2) dem Wesen nach mit denselben Methoden behandelt werden

kann, mit welchen Verfasser die absolute Konvergenz der Entwicklung nach Orthogonalpolynomen untersuchte. Unser Resultat kann in folgenden beiden Sätzen zusammengefaßt werden:

SATZ I. *Es sei* $(a, b) \subset [-1, +1]$ *und*

$$w(x) \geqq m > 0 \quad \textit{für} \quad x \in (a, b); \tag{3}$$

*ferner sei* $f(x)$ *eine in* $[-1, +1]$ *k-mal stetig differenzierbare Funktion, für welche*

$$\sum_{n=1}^{\infty} \frac{\omega\left(\frac{1}{n}, f^{(k)}\right)}{\sqrt{n}} < \infty \tag{4}$$

*besteht, wobei* $\omega(\delta, f^{(k)})$ *den Stetigkeitsmodul von* $f^{(k)}(x)$ *bedeutet; dann ist die Reihe* (2) *gleichmäßig in jedem inneren Teilintervall von* $(a, b)$ *absolut konvergent und stellt dort die Funktion* $f^{(k)}(x)$ *dar.*

SATZ II. *Es genüge* $w(x)$ *denselben Bedingungen, wie in Satz I und es sei außerdem* $w(x) \leqq W(1-x^2)^{-\frac{1}{2}}$. *Sei ferner* $f(\cos\vartheta) = g(\vartheta)$ *und bezeichne* $\omega_k^{(2)*}(\delta, f)$ *den Stetigkeitsmodul von* $\frac{d^k g}{d\vartheta^k}$ *bezüglich der Metrik* $L^2$, *also*

$$\omega_k^{(2)*}(\delta, f) = \left\{ \underset{|h| \leqq \delta}{\mathrm{Max}} \int_0^{\pi} [g^{(k)}(\vartheta + h) - g^{(k)}(\vartheta)]^2 d\vartheta \right\}^{\frac{1}{2}}, \tag{5}$$

*und es bestehe*

$$\sum_{n=1}^{\infty} \frac{\omega_k^{(2)*}\left(\frac{1}{n}, f\right)}{\sqrt{n}} < \infty, \tag{6}$$

*dann ist die Reihe* (2) *gleichmäßig in jedem inneren Teilintervall von* $(a, b)$ *absolut konvergent und stellt dort die Funktion* $f^{(k)}(x)$ *dar.*

Wir bemerken, daß (4) für $f^{(k)} \in \mathrm{Lip}\,\alpha$ mit $\alpha > \frac{1}{2}$ bestimmt befriedigt ist.

HILFSSATZ I. *Aus* (3) *folgt*

$$\sigma_n^{(k)}(x) \equiv \sum_{\nu=0}^{n} [p_\nu^{(k)}(x)]^2 = O(n^{2k+1}) \tag{7}$$

*gleichmäßig in* $x$ *für jedes innere Teilintervall von* $(a, b)$.

BEWEIS. $\Pi_n(x)$ durchlaufe alle Polynome höchstens $n$-ten Grades mit

$$\int_{-1}^{+1} [\Pi_n(t)]^2 w(t)\, dt \leqq 1. \tag{8}$$

Wir suchen das Maximum von $\Pi_n^{(k)}(x)$. Es ist offenbar

$$\Pi_n(t) \equiv \sum_{\nu=0}^{n} \gamma_\nu p_\nu(t)$$

mit

$$\sum_{\nu=0}^{n} \gamma_\nu^2 \leqq 1.$$

Also ist

$$\Pi_n^{(k)}(x) = \sum_{\nu=0}^{n} \gamma_\nu p_\nu^{(k)}(x) \leqq \left\{ \sum_{\nu=0}^{n} [p_\nu^{(k)}(x)]^2 \right\}^{\frac{1}{2}},$$

und das Gleichheitszeichen findet statt, wenn man

(9)
$$\Pi_n(t) = \left\{ \sum_{\nu=0}^{n} [p_\nu^{(k)}(x)]^2 \right\}^{-\frac{1}{2}} \sum_{\nu=0}^{n} p_\nu^{(k)}(x) p_\nu(t)$$

wählt.

Wir vergleichen nun $\sigma_n^{(k)}(x)$ mit dem für eine Gewichtsfunktion $\overline{w}(t) \leqq w(t)$ ähnlich gebauten Ausdruck $\overline{\sigma}_n^{(k)}(x)$. Dann ist die Bedingungsungleichung (8) für das Polynom (9) auch mit der Gewichtsfunktion $\overline{w}(t)$ befriedigt, und infolgedessen ist das entsprechende Maximum mit $\overline{w}$, das gleich $\{\overline{\sigma}_n^{(k)}\}^{\frac{1}{2}}$ ist, sicher nicht kleiner als $\{\sigma_n^{(k)}\}^{\frac{1}{2}}$. Es ist also

(10)
$$\sum_{\nu=0}^{n} [p_\nu^{(k)}(x)]^2 \leqq \sum_{\nu=0}^{n} [\overline{p}_\nu^{(k)}(x)]^2.$$

Es sei nun speziell

$$\overline{w}(t) = \begin{cases} m & \text{für } t \in (a, b), \\ 0 & \text{für } t \,\overline{\in}\, (a, b), \end{cases}$$

die entsprechenden normierten Orthogonalpolynome sind

$$\overline{p}_\nu(x) = \sqrt{\frac{(2\nu+1)m}{b-a}}\, P_\nu\left(-1 + 2\frac{x-a}{b-a}\right),$$

wobei $P_\nu(x)$, wie üblich, das $\nu$-te Legendresche Polynom bedeutet. Hieraus folgt $\overline{p}_\nu(x) = O(1)$ gleichmäßig in jedem inneren Teilintervall von $(a, b)$ und unter Anwendung des Satzes von BERNSTEIN

(11)
$$\overline{p}_\nu^{(k)}(x) = O(\nu^k)$$

ebenfalls gleichmäßig in jedem inneren Teilintervall von $(a, b)$. Aus (10) und (11) erhalten wir unmittelbar die Abschätzung (7), w. z. b. w.

Bezeichne nun

(12)
$$E_n^{(2)}(f; w) = \left\{ \sum_{\nu=n+1}^{\infty} c_\nu^2 \right\}^{\frac{1}{2}}.$$

HILFSSATZ II. *Die Gewichtsfunktion sei so beschaffen, daß die Derivierten der Orthogonalpolynome an der Stelle* $x$ (7) *befriedigen; ferner gelte für die Funktion* $f(x)$

$$\sum_{\nu=1}^{\infty} \nu^{k-\frac{1}{2}} E_\nu^{(2)}(f;w) < \infty; \tag{13}$$

*dann ist die Reihe* (2) *an der Stelle* $x$ *absolut konvergent.*

BEWEIS. Es ist, da $E_n^{(2)}(f;w)$ eine monoton abnehmende Funktion von $n$ ist,

$$\sum_{\nu=2^m}^{2^{m+1}-1} |c_\nu p_\nu^{(k)}(x)| \leqq \left\{ \sum_{\nu=2^m}^{2^{m+1}-1} c_\nu^2 \right\}^{\frac{1}{2}} \left\{ \sum_{\nu=0}^{2^{m+1}} [p_\nu^{(k)}(x)]^2 \right\}^{\frac{1}{2}} = E_{2^m}^{(2)}(f;w) \cdot O\left(2^{\left(k+\frac{1}{2}\right)m}\right) =$$

$$= O\left(2^{\left(k-\frac{1}{2}\right)m}\right) \sum_{\nu=2^{m-1}+1}^{2^m} E_n^{(2)}(f;w) = O\left\{ \sum_{\nu=2^{m-1}+1}^{2^m} \nu^{k-\frac{1}{2}} E_n^{(2)}(f;w) \right\},$$

und man addiere diese Ungleichung vom $m=1$ bis $m=\infty$.

HILFSSATZ III. *Es besteht die Ungleichung*

$$E_n^{(2)}(f;w) \leqq c_1 n^{-k} \omega\left(\frac{1}{n}; f^{(k)}\right), \tag{14}$$

*wobei die Zahl* $c_1$ *von* $n$ *unabhängig ist.*

BEWEIS. Bekanntlich ist

$$E_n^{(2)}(f;w) = \left\{ \operatorname{Min} \int_{-1}^{+1} [f(x) - \Pi_n(x)]^2 w(x)\, dx \right\}^{\frac{1}{2}},$$

wobei $\Pi_n(x)$ alle Polynome höchstens $n$-ten Grades durchläuft. Als Konkurrenzpolynom wählen wir das Polynom, welche die bestmögliche Approximation von $f(x)$ im Tschebyscheffschen Sinne liefert; die Abweichung sei $E_n(f;w)$. Nach dem Satze von JACKSON ist

$$E_n(f;w) \leqq c_2 n^{-k} \omega\left(\frac{1}{n}; f^{(k)}\right),$$

und somit besteht

$$E_n^{(2)}(f;w) \leqq \left\{ \int_{-1}^{+1} w(x)\, dx \right\}^{\frac{1}{2}} c_2 n^{-k} \omega\left(\frac{1}{n}; f^{(k)}\right),$$

w. z. b. w.

Aus den Hilfssätzen I, II und III folgt unmittelbar die absolute Konvergenz von (2). Es bleibt nur übrig zu beweisen, daß die Summe der in jedem inneren Teilintervall von $(a, b)$ gleichmäßig absolut konvergenten Reihe (2) die Funktion $f^{(k)}(x)$ darstellt.

Nach $k$-maliger Integration erhalten wir aber die orthogonale Polynomreihe (1), die ebenfalls in jedem inneren Teilintervall gleichmäßig absolut konvergent ist. Nach einem Satze des Verfassers [4] konvergieren die $(C, 1)$

Summen der Reihe (1) gegen $f(x)$, also ist auch die gewöhnliche Summe dieser Reihe gleich $f(x)$. Infolge der gleichmäßigen Konvergenz der differenzierten Reihe (2) ist die gliedweise Differentiation der Reihe (1) erlaubt, und hieraus folgt, daß die Summe der Reihe (2) gleich $f^{(k)}(x)$ sein muß. Um auch den Beweis des Satzes II zu Ende führen, muß Hilfssatz III durch den folgenden ersetzt werden:

HILFSSATZ IV. *Ist*

$$w(x) \leqq W(1-x^2)^{-\frac{1}{2}}, \tag{15}$$

*dann gilt*

$$E_n^{(2)}(f; w) \leqq c_3 n^{-k} \omega_k^{(2)*}\left(\frac{1}{n}; f\right). \tag{16}$$

Wir bezeichnen mit

$$a_n = \frac{2}{\pi}\int_0^\pi f(\cos\theta)\cos n\theta\, d\theta$$

die Koeffizienten der Cosinusentwicklung von $f(\cos\theta)$. Nach einem Lemma von G. ALEXITS [1] besteht für jede monoton nicht abnehmende Folge $\{\lambda_\nu\}$ positiver Zahlen

$$\sum_{\nu=0}^{\infty} \lambda_\nu c_\nu^2 \leqq \frac{W\pi}{2}\sum_{\nu=0}^{\infty}\lambda_\nu a_\nu^2. \tag{17}$$

Wir betrachten die Identität

$$\sum_{\nu=1}^{\infty} \nu^{2k} a_\nu^2 \left(1-\frac{\sin\frac{k}{n}}{\frac{k}{n}}\right) = \frac{n}{\pi}\int_0^{\frac{1}{n}}\int_0^{\pi}\left[g^{(k)}\left(\vartheta+\frac{h}{2}\right)-g^{(k)}\left(\vartheta-\frac{h}{2}\right)\right]^2 d\vartheta.$$

Die rechte Seite ist höchstens gleich $\omega_k^{(2)*}\left(\frac{1}{n}\right)$; beachten wir noch, daß für $k \geqq n$ der Faktor $1-\frac{\sin\frac{k}{n}}{\frac{k}{n}}$ oberhalb einer positiven Zahl $c_4^{-1}$ bleibt, so erhält man

$$\sum_{\nu=n+1}^{\infty} \nu^{2k} a_\nu^2 \leqq c_4\left[\omega_k^{(2)*}\left(\frac{1}{n}\right)\right]^2,$$

also wegen (17)

$$\sum_{\nu=n+1}^{\infty} c_\nu^2 \leqq n^{-2k}\sum_{\nu=n+1}^{\infty}\nu^{2k}c_\nu^2 \leqq \frac{W\pi}{2}n^{-2k}\sum_{\nu=n+1}^{\infty}\nu^{2k}a_\nu^2 \leqq \frac{W\pi}{2}c_4 n^{-2k}\left[\omega_k^{(2)*}\left(\frac{1}{n}\right)\right]^2$$

und damit haben wir (16) bewiesen.

Satz II folgt aus den Hilfssätzen I, II und IV.

*(Eingegangen am 18. März 1955.)*

## Literaturverzeichnis


[1] G. ALEXITS, Sur la convergence et la sommabilité presque partout des series de polynomes orthogonaux, *Acta Sci. Math.*, **12** (1950), S. 223—225.

[2] S. N. BERNSTEIN, Sur la convergence absolue des series trigonometriques, *Comptes Rendus Acad. Sci. Paris*, **199** (1934), S. 397—400.

[3] G. FREUD, Über die absolute Konvergenz von orthogonalen Polynomreihen, *Acta Math. Acad. Sci. Hung.*, **4** (1953), S. 127—135.

[4] G. FREUD, Über die starke $(C, 1)$-Summierbarkeit von orthogonalen Polynomreihen, *Acta Math. Acad. Sci. Hung.*, **3** (1952), S. 83—88.

[5] С. Б. Стечкин, Об абсолютной сходимости ортогональных рядов, Математический Сборник, **29** (1951), S. 225—232.


## О ПОЧЛЕННОМ ДИФФЕРЕНЦИРОВАНИИ РЯДОВ ПО ОРТОГОНАЛЬНЫМ МНОГОЧЛЕНАМ

Г. Фрайд (Будапешт)


(Резюме)

Пусть $w(x) \in L$ $(-1 \leqq x \leqq 1)$ — неотрицательная весовая функция; обозначим соответствующие нормированные ортогональные многочлены через $p_0(x), p_1(x), \ldots$.

Пусть

(1) $$f(x) \sim \sum c_\nu p_\nu(x)$$

разложения по системе $\{p_\nu(x)\}$ $k$-раз непрерывно дифференцируемой функций $f(x)$. $k$-кратным формальным дифференцированием отсюда получаем

(2) $$f^{(k)}(x) \sim \sum c_\nu p_\nu^{(k)}(x).$$

Имеет место следующая

Теорема. Пусть

(3) $$w(x) \geqq m > 0 \qquad (-1 \leqq a < x < b \leqq 1)$$

и, обозначая модуль непрерывности функции $f^{(k)}(x)$ через $\omega(\delta, f^{(k)})$, пусть

(4) $$\sum_{n=1}^{\infty} \frac{\omega\left(\frac{1}{n}, f^{(k)}\right)}{\sqrt{n}} < \infty;$$

тогда (2) во всяком внутреннем подинтервале $(a, b)$ сходится абсолютно и равномерно, и его сумма равно $f^{(k)}(x)$.

Условие (4) можно заменять следующим: Пусть

$$w(x) \leqq w(1-x^2)^{-1/2}$$

и для $g(\vartheta) = f(\cos\vartheta)$

$$\sum_{n=1}^{\infty} \frac{\omega^{(2)}\left(\frac{1}{n}, g^{(k)}\right)}{\sqrt{n}} < \infty.$$

Les Acta Mathematica paraissent en russe, français, anglais et allemand et publient des mémoires du domaine des sciences mathématiques.

Les Acta Mathematica sont publiés sous forme de fascicules qui seront réunis en un volume.

On est prié d'envoyer les manuscrits destinés à la rédaction et écrits à la machine à l'adresse suivante:

*Acta Mathematica, Budapest 62, Postafiók 440.*

Toute correspondance doit être envoyée à cette même adresse.

Le prix de l'abonnement est de 110 forints par volume.

On peut s'abonner à l'Entreprise pour le Commerce Extérieur des Livres et Journaux „Kultúra" (Budapest, VI., Sztálin út 21. Compte-courant No. 43-790-05-1817) ou à l'étranger chez tous les représentants ou dépositaires

---

The Acta Mathematica publish papers on mathematics in Russian, French, English and German.

The Acta Mathematica appear in parts of various size, making up one volume.

Manuscripts should be typed and addressed to:

*Acta Mathematica, Budapest 62, Postafiók 440.*

Correspondence with the editors and publishers should be sent to the same address.

The rate of subscription to the Acta Mathematica is 110 forints a volume. Orders may be placed with „Kultúra" Foreign Trade Company for Books and Newspapers (Budapest, VI., Sztálin út 21. Account No. 43-790-057-181) or with representatives abroad.

---

Die Acta Mathematica veröffentlichen Abhandlungen aus dem Bereiche der mathematischen Wissenschaften in russischer, französicher, englischer und deutscher Sprache

Die Acta Mathematica erscheinen in Heften wechselnden Umfanges. Mehrere Hefte bilden einen Band.

Die zur Veröffentlichung bestimmten Manuskripte sind, mit Maschine geschrieben, an folgende Adresse zu senden:

*Acta Mathematica, Budapest 62, Postafiók 440.*

An die gleiche Anschrift ist auch jede für die Redaktion und den Verlag bestimmte Korrespondenz zu richten.

Abonnementspreis pro Band 110 Forints Bestellbar bei dem Buch- und Zeitungs-Außenhandels-Unternehmen „Kultúra" (Budapest, VI., Sztálin út 21. Bankkonto Nr. 43-790-057-181) oder bei seinen Auslandsvertretungen und Kommissionären.

56,— Ft

INDEX



Technikai szerkesztő: Molnár Ferenc

A kiadásért felelős: az Akadémiai Kiadó igazgatója — Műszaki felelős: Farkas Sándor

A kézirat beérkezett: 1955. IV. 12. — Terjedelem: 19½ (A/5) ív, 2 ábra.

Szegedi Nyomda Vállalat, Szeged, 55-1876 — Felelős vezető: Vincze György